UNDERSTANDING STREAMFLOW CHANGE AT THE SCALE OF A MAJOR CITY: CHICAGO

BY

CHARLES JACQUES ROUGÉ

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Professor Ximing Cai

Abstract

Urban and economic development and climate change have made the non-stationarity of streamflow records an emerging phenomenon in hydrology. This thesis studies the systematic shifts in streamflow that have been observed in urban, suburban and agricultural watersheds in and around the Greater Chicago area. A novel statistical method, the Procedure for Change Pattern and Significance (PCPS), is developed to detect non-stationarity and identify the timing and duration of these shifts in hydrologic records. This method uses the rank correlations of the Mann-Kendall statistics to extract information that can be interpreted to identify the timing and duration of a detected shift, while it also adopts the Mann-Kendall test to determine whether the change is statistically significant. It is also shown that incorporating the aspects of timing, duration and significance in a consistent analytical framework enables the issue of serial correlation to be tackled better. The proposed statistical framework is associated with conceptual modeling tools to understand how adaptive human responses to urbanization mitigate and sometimes even offset the effects of land-use change in a large urban area, such as the Greater Chicago area. Results of PCPS, along with an analysis of the recession process, show that while impervious surfaces increase flooding, stormwater management facilities mitigate or even counter this impact over a wide range of scales. On the other hand, an urban water balance shows that low flows have increased since the impacts of effluent discharges, pipe leakage and garden irrigation from municipal water systems using water from Lake Michigan outweigh the reduction of base flow due to impervious surfaces. In addition, the increasing trend of the mean flow in the area can be explained by the combined impacts of water withdrawals from Lake Michigan, land-use change and climatic variability. Finally, it is found that a step increase in rainfall coincides with a step increases in streamflow around 1970. Thus, PCPS detects the impacts of climate and urbanization on streamflow throughout the Greater Chicago area.

To my whole family.

A toute ma famille.

Acknowledgments

This project would not have been possible without the support of many people. My first thanks go to my adviser, Dr. Ximing Cai, his invaluable insights and thurough support. Thanks also to the students from Dr. Cai's group, for their feedbacks and advice, with special thoughts to Dr. Dingbao Wang, Dr. Mohamad Hejazi and Jory Hecht. Further thanks to Dr. Cai and the Department of Civil and Environmental Engineering for their continued financial support, without which this work could never have been carried out.

Finally, because academic work can seldom be accomplished without a sound social life, I would like to thank my friends for their friendship, and my family for their love.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis deals with the issue of non-stationarity, one that has been at the forefront of discussion in hydrological sciences since the benchmark paper from *Milly et al.* (2008). Its focus is to carry out a statistical analysis of past streamflow and rainfall records to understand the impacts of urban development and climatic variability at different scales in the greater metropolitan area of a major city, Chicago. Three major issues are addressed in this thesis to advance the understanding of hydrologic non-stationarity:

1. Understanding the causes of non-stationarity and whether they are related to climate change, local urban development, or both.

2. Understanding the spatial and temporal scales at which these causes come into play, which is a prerequisite for understanding their real impact on the hydrologic cycle.

3. Developing a new method for a multi-scale analysis of the causes of non-stationarity.

Stationarity has long been a common and convenient assumption in water resources planning and management. Under it, simple methods can be used to extract from the data all the useful hydrologic indicators and to provide estimates that can be refined year after year as the records become longer. For example, the 100-year flood level of a river is a variable that can be computed from the series of the annual maxima of the flow, using simple methods that can be found in any classic hydrology textbook, such as *Chow, Maidment, and Mays* (1988). This knowledge about 100-year flood levels is then used to build stormwater management facilities, such as culverts and storm sewers.

However, since the forcing of human activity on the hydrologic cycle has become more and more apparent at a wide range of spatial and temporal scales, experts have come to question this assumption. As *Milly et al.* (2008) note, different aspects of human development, including land-use and land cover change, have long compromised the assumption of stationarity at the basin level. Most importantly, anthropogenic climate change is bound to significantly alter the hydrologic cycle at a global scale, and its impact on streamflow has already been detected by some studies (e.g *Barnett et al.*, 2005; *Milly et al.*, 2005). They therefore conclude that "Stationarity is dead [...], and cannot be revived". They eventually call for innovative thinking and methods to provide estimates of hydrologic indicators that would be both reliable and useful for water management.

Unfortunately, the detection of such shifts in streamflow is complicated by the non-linearity and the complexity of the diverse phenomena coming into play in the hydrologic cycle. One can for instance cite the existence of long-range memory in streamflow series (*Pelletier and Turcotte*, 1997; *Koutsoyiannis*, 2002, 2003; *Montanari*, 2003; *Yue and Gan*, 2004; *Koscielny-Bunde et al.*, 2006) and the tremendous natural variability that these stochastic processes incur (*Koutsoyiannis*, 2006). They can ultimately be linked with the high-order non-linearity of the hydrologic system, which is sometimes termed chaos (*Sivakumar*, 2008). Thus, many hydrologic processes display threshold responses to human and natural forcings, for which a physical understanding is necessary to predict them (*Zehe and Sivapalan*, 2009). Meanwhile, the modeling of non-linearity and threshold behaviors in environmental systems faces challenges in terms of physical understanding (e.g. *Scheffer et al.*, 2001) and mathematical modeling (e.g. *Martin*, 2004).

A step in the direction of understanding non-stationarity is to be able to trace it back to its causes, even in the most complex situations. Indeed, in many cases, both climate and land-use change can impact the hydrological cycle simultaneously (e.g. *Hejazi and Moglen*, 2007, 2008). Therefore, they should be studied jointly (*Claessens et al.*, 2006). To investigate hydrologic non-stationarity, we can use past examples of how these two types of factors can alter water resources. These include the dwindling Aral Sea levels, which are falling mainly due to the development of unsustainable irrigation as well as a long-term decrease in rainfall (e.g. *Glantz*, 1999). Water resources in the Yellow River basin in China also faced the joint problem of dwindling rainfall and a growing and mismanaged need for water, which has caused its lower reaches to dry up (e.g. *Fu et al.*, 2004; *Cong et al.*, 2009). In both cases, gradual long-term evolutions interact with abrupt step changes in what can be considered a coupled human and natural system (*Liu et al.*, 2007a) in which feedbacks and interactions are complex and occur at different spatial and temporal scales (*Liu et al.*, 2007b). Thus rainfall variability in the Yellow River basin can be attributed to both gradual long-term trends (*Fu et al.*, 2004; *Cong et al.*, 2009) and step changes in its headwaters (*Zheng et al.*, 2007). Statistically detecting the effects of climate change alone on hydrology remains a challenge in many situations (*Kundzewicz et al.*, 2005; *Svensson et al.*, 2005; *Villarini et al.*, 2009a).

Yet it is important for planning and management purposes to understand the respective contributions of both climate variability and human interferences to streamflow, especially in urbanizing areas. Many recent studies have aimed to disentangle the impacts of climate change and urban development (e.g. *Hejazi and Markus*, 2009; *McCormick et al.*, 2009; *Villarini et al.*, 2009b). These studies have a crucial societal relevance because with more than half of the world population now living in cities (*Grimm et al.*, 2002), understanding how the growth of urban centers affects the hydrologic cycle is crucial for a better management of present and future water resources. Indeed, urban population growth is believed to become the major driver of water stress worldwide by 2025 (*Vörösmarty et al.*, 2000).

Unfortunately, many of these urban development analyses fail to take into consideration different spatial scales, including that of a whole metropolitan area. In a coupled human-nature system as complex as an urbanizing environment, multi-scale studies are needed to understand the mechanisms that can cause hydrologic change (*Liu et al.*, 2007b). It has recently been acknowledged that the consideration of different temporal scales was critical for understanding these mechanisms (*Claessens et al.*, 2006). The impacts of land-use change and urban development at different levels of spatial organization, from a household to a metropolitan area, has been explored through agent-based models (e.g. *Evans and Kelley*, 2004; *Bolte et al.*, 2006; *Monticino et al.*, 2007), but studies of their specific effects on the hydrologic cycle have typically been conducted at only one spatial scale.

These studies were useful for determining the existence of human interferences on streamflow. Stream channel alteration, interbasin transfers, leakage, effluent discharge and pumping can impact base flow and low flows (*Barringer et al.*, 1994; *Meyer*, 2005; *Claessens et al.*, 2006; *Wang and Cai*, 2009) while stormwater detention basins can mitigate flood events (e.g. *Solo-Gabriele and Perkins*, 1997; *Yeh and Labadie*, 1997). These direct modifications to streamflow are distinct from the indirect effects of land-use change on streamflow, which have long been documented (e.g. *Hollis*, 1975; *Ferguson and Suckling*, 1990; *Dow and DeWalle*, 2000). In fact, these direct modifications have mitigated or even offset the effects of land-use change in certain situations (e.g. *Barringer et al.*, 1994; *Burns et al.*, 2005).

In a context of such complex interactions, understanding the spatial scales associated with a particular type of human interferences is necessary for understanding the final effects of urban development on the hydrologic cycle. Indeed, evaluating the impact of a human-induced mechanism assumes knowledge of the scales at which impacts occur (*Liu et al.*, 2007b). This knowledge would also motivate the search for the same human interferences in other places. In its absence, only the well-known effects of land use change can be taken into account when exploring the effects of climate and urban development in simulation models (e.g. *DeWalle et al.*, 2000; *Hejazi and Moglen*, 2007, 2008; *Hurkmans et al.*, 2009).

This study thus has to simultaneously consider numerous human interferences on streamflow at different spatial and temporal scales, but also the impact of climate fluctuations at these scales, a complicating factor when studying human interferences on the alteration of hydrologic processes such as streamflow (*Claessens et al.*, 2006; *McCormick et al.*, 2009). Moreover, while the comparison of urban and rural basins provides a good means of investigating the effects of urban development (*Lazaro*, 1976; *Changnon and Demissie*, 1996), agricultural areas in Northeastern Illinois also have experienced land-use change through drainage tiles (*Changnon and Demissie*, 1996). These intertwined relationships make the proposed statistical assessment very complex. To deal with the presence of different temporal scales of change, we need statistical methods to help us detect the varied change patterns present in the data by not only documenting the existence and significance of change, but also describing its temporal location and duration. To date, most studies have focused on the mere detection of significant streamflow change itself (*Lins and Slack*, 1999; *Douglas et al.*, 2000; *Kundzewicz et al.*, 2005; *Svensson et al.*, 2005) while only a few studies (e.g. *McCabe and Wolock*, 2002; *Villarini et al.*, 2009a) have aimed to detect the pattern of change. Investigation by statistical analysis alone is usually considered to be insufficient for understanding the causes of change, since only process-based models can achieve this goal (*Lettenmaier et al.*, 1994; *Groisman et al.*, 2004; *Kumar et al.*, 2009).

In fact, instead of learning about the change patterns through data analysis, it is recommended to know which one is present prior to starting the analysis (*Kundzewicz and Robson*, 2004). Even cutting-edge time-series modeling techniques like GAMLSS (Generalized Additive Models for Location, Scale and Shape) require the modeler to first look for step changes (*Villarini et al.*, 2009a), or to first assume gradual changes, e.g. exploring effects of urbanization and climate variability in the Charlotte, North Carolina area (*Villarini et al.*, 2009b). This disregards the fact that if a given cause for change acts on a system for a long time, it does necessarily imply the effects are going to be felt gradually, as nonlinearities often allow many sudden and unexpected shifts to occur in natural systems disturbed by gradual forcings (*Scheffer et al.*, 2001).

We thus need a method that allows for meaningful interpretations on change patterns to be made directly from the data. This method must also be non-parametric, because parametric methods require the assumption of normality, which is often invalid in hydrology (*Villarini et al.*, 2009a). Recall the main objective of this thesis is to understand hydrologic non-stationarity in the Chicago metropolitan area by retrieving relevant information from the data (issues 1 and 2). To address this objective, it is necessary to develop a novel method (issue 3)). Then this methodological development can be incorporated it into an integrated framework dealing with timing, duration and significance of changes.

This thesis is organized as follows:

Chapter 2 explores the capabilities of trend and change-point tests to derive a method for providing insights on the change pattern in a given time-series. This method, called PCPS (Procedure for Change Pattern and Significance) is then integrated into a holistic procedure that assesses the change pattern and determines its significance, and addresses technical issues, such as serial correlation. The limitations of this method, as well as its applicability for regional streamflow assessment, are also discussed.

Chapter 3 presents the hydrologic data used to detect change in northeastern Illinois and the results of the application of the method developed in Chapter 2 in order to tackle issues 1) and 2). It will depict the main effects of climate on these changes. It will use ten gauges in agricultural basins around Chicago, with a different history of land-use change. These gauges serve as a means of understanding the impact of rainfall variability on streamflow without the confounding effect of urbanization. Then the impacts of urbanization across various spatial and temporal scales will be described, with a focus on the effects that it induces on streamflow in the whole Chicago metropolitan area. All aspects of streamflow, from low flows to annual maxima, will be discussed using long-term gauging stations with contributing areas spanning a wide range of spatial scales, ranging from the smallest urban catchments (around twenty square kilometers) much larger ones that cover a considerable portion of the Greater Chicago area. Other climatic data, along with conceptual models, will be used or derived to be compared to the results of the PCPS method and clarify their physical significance.

Finally, Chapter 4 contains the conclusion to this thesis. It summarizes the findings of Chapters 2 and 3, and highlights the insights that the PCPS method can bring into the analysis through the application to streamflow change in the Greater Chicago area. Finally, it presents the limitations of this analysis and identifies future research directions that could overcome these limitations.

Chapter 2

Understanding change in a time-series

2.1 Introduction

In hydrologic time-series, statistical tests are designed to assess the significance of change at the desired significance level (*Kundzewicz and Robson*, 2004). In particular, trend tests determine whether there is change based on the behavior of the whole series. They have been extensively used to detect change in hydrological records, such as streamflow (e.g. *Lins and Slack*, 1999, 2005; *Zhang et al.*, 2001; *Kundzewicz et al.*, 2005; *Svensson et al.*, 2005), rainfall (e.g. *Karl and Knight*, 1998; *Pryor et al.*, 2009) or both streamflow and rainfall (e.g. *Lettenmaier et al.*, 1994; *Groisman et al.*, 2001; *Small et al.*, 2006). In such studies, one can then identify the spatial pattern of change by testing numerous time-series in a given area.

Another issue is to obtain the temporal pattern of change, which can be defined by its timing and duration, i.e. when does change occur and how long does it last? As highlighted in Chapter 1, understanding these patterns can be crucial in the emerging context of non-stationarity. Hence, the goal of this chapter is to design a comprehensive framework that systematically addresses the issues of significance, timing and duration of change. That would enable to understand the spatio-temporal pattern of change in a given situation, so as to describe the effects of change before using other tools to understand the causes. This introductory section is to review how existing methods are dealing with the issues of timing and significance, so as to set the theoretical background for this work.

Trend detection methods are not designed to investigate the timing or the duration of change. Some studies use several different temporal windows to obtain an improved understanding of the timing of change (e.g. *Douglas et al.*, 2000; *Zhang et al.*, 2001) but that often translates into comparing periods of different lengths, while the length of a time-series has an influence on the outcome of statistical tests.

To determine the timing, change-point tests can be performed in association with a trend test, whether it be by a bayesian approach (*Xiong and Guo*, 2004; *Zhang et al.*, 2009), a moving t-test (*Zheng et al.*, 2007) or the nonparametric Pettitt test (*Tomozeiu et al.*, 2000). The trend test is performed first, to diagnose whether there is statistically significant change in the entire time series. Then, the change point test is to identify the year(s) that contain the most information about change. Such assessments then enable to look for related events that occur around the same times. Change-point tests can also be performed alone to provide information on both the significance and timing of change (e.g. *Lazaro*, 1976; *Buishand*, 1982; *Perreault et al.*, 1999; *Yue and Wang*, 2002a,b). But the notion of duration is absent when change-point tests are used. It cannot tell whether the observed shift is abrupt (a step change) or gradual (a trend).

In fact, when dealing with trends or steps it is recommended that one should rather use a trend test to deal with a gradual change, or a change-point test to deal with a shift (*Kundzewicz and Robson*, 2004). Due to the association of a type of tests with the pattern it is supposed to detect, *Yue et al.* (2002a) exclusively use series with a linear trend to compare the power of two trend tests. Likewise, *Yue and Wang* (2002a) exclusively use series with a step change to test the power of the Mann-Whitney change-point test.

But while assuming a given change pattern sometimes determines the way methods are used, detection of change using a given method may in turn lead to assuming what its temporal pattern is. Thus, *McCabe and Wolock* (2002) note that when it comes to trend tests, "a statistically significant result generally is interpreted to indicate a monotonic trend"; they rightfully add that "previous research has not demonstrated, however, whether commonly used statistical tests for trends can distinguish a gradual monotonic change from an abrupt "step" change". By performing a trend test on a moving temporal window, they detect a step change in streamflow in the conterminous United States around 1970. A recent study by *Villarini et al.* (2009a) further questions the assumption that a statistically significant trend test is an indication of the existence of a gradual change. It puts forward a method that tries to successively detect sudden and gradual changes in a same time-series. Only few trends are found after the diagnosed change points. But the example of the Yangtze River (*Xiong and Guo*, 2004) shows that a trend test can lose its significance after removing the change point, and conversely.

As both *McCabe and Wolock* (2002) and *Villarini et al.* (2009a) also perceive, the issue of deciding between gradual trend and step change is linked to knowing whether change may continue after the end of the time-series or not. Indeed a gradual trend is often associated to a change that continues after the time-series ends, even though this assumption may be misleading (*Koutsoyiannis*, 2006). A step change, on the other hand, reflects historical change. This chapter will propose ways to address the issue of the statistical detection of current change, by applying to it the novel developments in pattern recognition.

In fact, both works by *McCabe and Wolock* (2002) and *Villarini et al.* (2009a) use trend tests to extract more from them than what they were originally designed for, that is, to know whether change is significant. But both methods rely on splitting the time-series in a clever way, then still use test significance as a clue to detect pattern. The present work goes beyond that conception when using the Mann-Kendall trend test (*Mann*, 1945; *Kendall*, 1975), which is widely recognized as one of the most popular trend tests in the field (*Kundzewicz and Robson*, 2004; *Villarini et al.*, 2009a). Like in *Villarini et al.* (2009a), this trend test is related to a suitable change-point test. It is aimed at building a comprehensive framework that deals with steps and gradual changes alike and without prior assumption on the change that may be present in the data. The conceptual framework is represented on Figure 2.1. While tests are still used in this thesis to detect the significance of change, the main innovation is to directly use the test statistics to extract information about the change pattern in terms of timing and duration.

One last issue a framework that addresses the significance of change must tackle is that of correlation. Dealing with correlation in hydrologic data has been a constant concern since the study of water quality data by *Lettenmaier* (1976), giving rise to a variety of modifications to statistical tests. In the case of the MK trend test, these can account for seasonality (*Hirsch and Slack*, 1984) or spatial correlation (*Douglas et al.*, 2000). Here we are going to deal only with serial correlation, and combine the lessons of different previous studies.

The chapter is organized as follows. Section 2.2 presents the background of statistical methods that are both already available and used in this thesis. Then in section 2.3 we will explore how available non-parametric tests can identify a change pattern in a time-series; the insights gained will be used to derive a method describing the timing and duration of change in section 2.4. The method is to be then incorporated to an integrated change assessment procedure that simultaneously deals with the two aspects of a change pattern, i.e. timing and duration, as well as its statistical significance. As such, it also tackles the common issue of serial correlation. Furthermore, thoughts will be presented to tackle the important issue of examining if current change can be chosen as an acceptable model. The applicability and limits of the proposed holistic methodology will be discussed in section 2.5, based on hypothetical examples designed to better understand the insights it can bring. A framework for a thorough and rigorous validation will be outlined in section 2.6. The concluding section 2.7 will then summarize the findings.

This chapter also features numerous Monte-Carlo experiments. All are performed with 10,000 runs and normal time-series, long of 100 points (unless stated otherwise).

2.2 Statistical tests

This section is to review the statistical tests used in this chapter, and in this thesis in general. We will focus on tests that are non-parametric, like the method presented later on. In many hydrological time-series, the assumption of normality that is usually needed to use parametric methods breaks down. Such series are usually skewed, and in the case of peak flow series, can contain outliers that are likely to bias parametric estimates. Non-parametric methods are only based on the rank of a measurement in the sample, which does not depend on the distribution of the stochastic process at its origin. As such, these methods are also called distribution-free.

We are here going to present a trend test and a change-point test. An estimator of slope is also used. Eventually, resampling method that can be used as a replacement for the Mann-Kendall test will be introduced. For a more thorough introduction on these methods, one can refer to

(*Kundzewicz and Robson*, 2004) and to (*Villarini et al.*, 2009a).

It is worthwhile at this point to introduce some general terminology from the fields of statistics. A statistics is a quantity computed from the sample of interest, and is then used to determine whether the basic hypothesis (called null hypothesis) of a test is true. There are two types of errors for a statistical test: the type I error occurs when the null hypothesis is rejected while it is true: it corresponds to the significance level α of the test. Conversely, the type II error occurs when the null hypothesis is rejected while it should not. For simplicity, here we will call power of a test the rejection rate of the null hypothesis; which is slightly different from the usual definition, where the power is only defined as the probability of rejecting the null hypothesis when it is not true (e.g. *Yue and Wang*, 2002a; *Yue et al.*, 2002a).

2.2.1 The Mann-Kendall trend test

The two majors non-parametric trend tests used in hydrology are the Mann-Kendall (MK) test (*Mann*, 1945; *Kendall*, 1975) and Spearman's rho (see for example *Lehmann*, 1975). They are found to have the same power by *Yue et al.* (2002a). Both are also based on the same assumptions and on the same null hypothesis. MK is chosen to detect change patterns in this thesis (section 2.4).

It is aimed at testing the null hypothesis H_0 that the observations are independent with no systematic trend. On a time-series of n observations, the MK statistics uses the sign (sgn) of the difference of two distinct observations i and j, defined as:

$$\operatorname{sgn}(x_j - x_i) = \begin{cases} 1 & \text{if} \quad x_j > x_i \\ 0 & \text{if} \quad x_j = x_i \\ -1 & \text{if} \quad x_j < x_i \end{cases} \tag{2.1}$$

The MK statistics can then be computed as follows:

$$S = \sum_{1 \leq i < j \leq n} \operatorname{sgn}(x_j - x_i) \tag{2.2}$$

so that S compares the total number of pairs that correlate the passing of time with an increase, to those that correlate it with a decrease. For $n \geq 8$, S is approximately normally distributed with mean and variance expressed as:

$$E(S) = 0 \tag{2.3}$$

$$V(S) = \frac{1}{18}\left(n(n-1)(2n+5) - \sum_m t_m m(m-1)(2m+5)\right) \tag{2.4}$$

where t_m is the number of ties of extent m (there is a tie when $x_j = x_i$, and the number of realizations that have the same value give the extent of the tie). In practice, using the MK test is not recommended for time-series with many ties. Now supposing there is no serial correlation in the data, H_0 is rejected at the level of significance α when:

$$S < z_{\frac{\alpha}{2}}\sqrt{V(S)} \quad \text{or} \quad S > z_{1-\frac{\alpha}{2}}\sqrt{V(S)} \tag{2.5}$$

where z_q is the q^{th} quantile of the standard normal distribution.

Taking the above equation along with equation (2.4) shows that significance at any level α is roughly proportional to $n^{3/2}$, while from equation (2.2) the number of terms in S increases like n^2 (it is equal to $n(n-1)/2$). As a consequence, it is easier to detect trends from longer records. This is consistent with the Monte-Carlo simulations by *Yue et al.* (2002a).

2.2.2 The Pettitt change point test

As *Kundzewicz and Robson* (2004) state, the Pettitt test (*Pettitt*, 1979) can be preferred over alternative non-parametric tests like Mann-Whitney because it is more powerful and more robust to changes in variance.

Similarly to the MK test, the Pettitt test assumes that that the observations are independent. It tests the null hypothesis H_0 that, when splitting the samples in two, there is no change in the median. It produces a rank-based comparison between the observations situated before and after a date t, that we can note $k(t)$:

$$k(t) = \sum_{i=1}^{t} \sum_{j=t+1}^{n} \operatorname{sgn}(x_j - x_i) \tag{2.6}$$

It then returns the time where the amount of information gathered by what we will call from now on the Pettitt point statistics $k(t)$ is the greatest. Noting T this time and K this amount of information, we have:

$$T = \arg\max(k(t)) \tag{2.7}$$

$$K = \max_{1 \leq t \leq n} (|k(t)|) \tag{2.8}$$

K is the final Pettitt statistics, and T will be called “Pettitt point” from now on. The significance probability associated with the rejection of H_0 is approximated by:

$$p \approx 2 \exp\left(\frac{-6K^2}{n^3 + n^2}\right) \tag{2.9}$$

with an accuracy within 1% for $p \leq 0.5$. Like for the MK test, the statistics K depends quadratically on the sample size n, so that equation (2.9) compares a quantity that is proportional to n^4 to another that is proportional to n^3. Once again, the shorter the time-series, the more difficult it is to detect a statistically significant change.

However, it is important to note that there is no confidence level associated with the date of change T. Using the Pettitt test, *Tomozeiu et al.* (2000) or *Villarini et al.* (2009a) explain that the underlying cause for the date of change has to be explained, in order to make it accepted as meaningful. Yet, we want to ensure its ability to locate change, as a modified version of the Pettitt statistics will be used in that purpose in section 2.4. That is why we compare it to an alternative parametric method such as ordered clustering, used in studies on hydrologic non-stationarity like *Wang et al.* (2009). This particular method minimizes the RMSE while splitting the series in two periods over which it fits distinct trends, and also finds the break point that achieves the best result.

Figure 2.2 shows the ability of the Pettitt test to locate a change point in a series with a step change. It displays the results of Monte-Carlo simulations on standard normal time-series over which a step change is superimposed at $t = 20$, 30, 40 and 50. While the ability of the Pettitt test to detect change decreases faster than that of the ordered clustering method, in this example of a step change it remains better than its counterpart. Knowing that parametric methods are often more powerful than non-parametric ones when the normality assumption holds (*Good*, 2005), this can ensure the capacity of the Pettitt test to locate the most relevant change point.

2.2.3 Sen-Theil estimate of the slope

When trends are assessed, slope calculations are carried out with the estimate developed by *Theil* (1950) and *Sen* (1968). Considering observations $(x_1, \ldots, x_n)$ at times $(1, \ldots, n)$, the Sen-Theil estimate is:

$$\beta = \text{median}_{1 \leq i < j \leq n} \left[\frac{x_i - x_j}{i - j} \right] \tag{2.10}$$

The above equation means that β is based on the ranks of the different possible estimates of the slope. So, it is a non-parametric indicator, with all the advantages we discussed earlier. This makes it a better alternative to parametric methods like least square estimators or linear regression methods.

Confidence intervals are associated with the estimator provided by equation (2.10). For example, if the lower 5% significance level is positive, the confidence level for the existence of an upwards trend is 95%. However, this can only be used for changes that can be assumed to be gradual and monotonic, like runoff timing changes due to a progressively shorter winter in western Canada (*Déry et al.*, 2009), or the mechanisms that led to water stress in the Yellow River basin (*Cong et al.*, 2009). Moreover, the magnitude of the confidence intervals increases tremendously as the sample size decreases, so that similar to what happens with the MK test, it becomes very difficult for the trend estimate to be significant for small sample sizes. As we are dealing with short (under 100 points) time-series with possibly very different change patterns rather than with monotonic, gradual ones, we cannot use the Sen-Theil estimate to assess the significance of a trend.

2.2.4 A resampling method as a trend test

Sometimes the MK test cannot be used because there are too many ties in the data, or because we are using subsets of time-series that are potentially too short. In such cases equation (2.4) is not relevant any more. But equation (2.2) still yields a symmetric (non-skewed) statistics. In such cases, according to *Good* (2005), permutations tests are exact. The procedure applied in this work is as follows: producing $1,000$ series from permutations of the original series and comparing their MK statistics to the original one. For the studied series to exhibit change at the 5% level, its MK statistics should be among the 25 largest or smallest, depending on whether it documents an increase or a decrease.

Like for the original MK test, this trend detection method is sensitive to serial correlation in the data. Indeed permutations break the original serial dependence, thus comparing data with a definite correlation structure to other data that do not have it. Such a comparison is irrelevant because, as we shall see later on in section 2.4.2, the variance of the MK statistics changes with the correlation structure.

2.3 Challenges in identifying patterns through statistical tests

This section only uses the two most basic types of shift, which are also the most commonly used to describe change: linear trends and step changes. They are also the two extremes in term of duration of change. This section aims at exploring how these two patterns can be distinguished, which is a prior to detecting shifts of different durations. In section 2.3.1 the limits of using current methods to recognize them will be discussed, then in section 2.3.2 the difficulty of extracting even these two pattern will be addressed. These limits of current methods justify an innovative use of the tests' statistics to differentiate linear trend from step change (section 2.3.3).

We will deal with series of 100 points because the power of all these tests increases with the length of the series, and because hydrological records of 100 points can be considered as long when one takes one point a year (*Villarini et al.*, 2009a).

2.3.1 The difficulty of change pattern recognition

To understand the limits of using test significance to detect change patterns, let us examine here the latest development for change pattern recognition in the field of hydrology (*Villarini et al.*, 2009a), using the example put forward by these authors. They base it on the argument that most of the time, step changes are overlooked because trend tests are performed first. Indeed, if the change is statistically significant, they argue that the common conclusion is then that the time series under study is undergoing a monotonic trend. As such, change is expected to still occur in the future, with all the consequences that can have for planning and management. To support their claim, the authors show the case of a series such as the one presented on the top left-hand corner of Figure 2.3. It has 100 normal realizations that have a standard deviation of 2. For the first 50 ones, the mean is 10 while for the last 50 ones the mean is 12. A linear trend can be fitted to the data but can be a misleading model. As a consequence, the recommended methodology is to perform a step change test first, then a trend test should be applied before and after the alleged change point.

Yet one can wonder whether conversely, searching for abrupt shifts first may lead to overlooking gradual trends, even though these are present. To test that, we apply the method proposed by *Villarini et al.* (2009a), and use Monte-Carlo experiments ($10,000$ runs) to compare how it distinguishes between a series like the one they propose and one with a linear trend. This latter type of series should display the same length and magnitude of change as the former one. Magnitude is measured here by the Sen-Theil estimate of the slope. The mean (over $10,000$ runs) of the series with a step change is 0.03, so that the series with a linear trend is built with a slope of 0.03 superimposed on 100 normal realizations with (top right Figure 2.3).

The results are presented in the bottom part of Figure 2.3, with a 5% significance level. We can check that both the Pettitt and MK tests can detect change in both types of time-series. And here, the change-point test is not the best at detecting point changes and the trend test is not the best at detecting trends. This counter-example poses both a terminology problem and a methodology problem, because it goes against the common association between test type and duration of change.

Yet, the really interesting result from Figure 2.3 is that trend detection before and after the identified change point yields a power of about 5% in both cases. This corresponds to the probability of a type I error occurring. It means the power of the MK test to detect a trend in the present case becomes almost nonexistent given the short length of record and the weak trend. Recall (*Yue et al.*, 2002a) that the power of the MK test depends on the slope β (indicative of the magnitude of change), the standard deviation σ of the process (indicative of its variability) and the length of record n. In the present case it becomes almost nonexistent given the short length of record and the weak trend. This example suggests that a method that relies only on the significance of tests to detect change patterns may be misleading. Here for instance, a step change would be detected regardless the real duration of change.

A practical consequence is that if an underlying mechanism for a step change exists in data, all other change patterns will be left out in many cases. Conversely, if we can explain a gradual change, why should we be bothered with explaining more local change patterns? In fact, both trend tests and change-point tests are not originally designed to recognize whether change can be explained by an abrupt step or something more gradual. And they cannot tell whether distinct patterns interact to make the change significant. This may prove unsatisfactory when simulations studies show that joint effects of climate and land-use change could induce more significant change in the hydrologic cycle than either one taken alone, for example in the Maryland Piedmont (*Hejazi and Moglen*, 2007, 2008).

2.3.2 The difficulty of change pattern separation

In the previous section we saw that there were limits for current methods to decide between trend and step. This section is to explore further the question of how we can separate them.

We use series where the gradual and step changes are superimposed. A first and immediate remark on the series superposing trend and step is provided by checking that contrary to what happens in Figure 2.3, the power of both the Pettitt and MK tests for the whole series is now 1 at the 5% significance level. In fact the blue line in Figure 2.4 shows that the power the Pettitt test is 1 even for a 99% confidence level. As expected, superposing two patterns of change increases its significance. This is a further indication that looking to describe change using a single pattern overlooks the contribution of all the other possible patterns present in the data.

As a consequence, one could want to be able to extract one pattern from the data without damaging the contribution of other patterns. As mentioned earlier, *Xiong and Guo* (2004) suggest that removing a trend can render the step non-significant, and conversely. But it may have been that only trend and step change taken together rendered the change pattern significant. An analysis of the problem via Monte-Carlo experiments is proposed now to know whether removing the shift in the data biases the trend assessment, and conversely.

First, let us remove the step change and compare the linear slope after removal to the one we have when only a trend exists ($\beta = 0.03$). The slope after removal is computed as the pondered average of the slopes computed before and after the Pettitt change point. If the shift removal were a neutral operation, then on ten thousand runs we should find an average of $\beta = 0.03$ after removal. In fact we find a normal distribution for β after removal, with parameters $\mu = 0.0087$ and $\sigma = 0.017$. In other words, on average almost 70% of the slope of the trend is getting removed that way. In fact, in more than 90% of the cases, a portion of the trend is removed.

Second, let us remove the linear trend we have in the data. Of course, the step change increases β, artificially increasing the magnitude of the linear trend. However, the detrended series still contains a centered shift in the median, so that the Pettitt test may still be able to detect change. In fact Figure 2.4 shows that the power of the Pettitt test at all levels above the 80% confidence level is much smaller than the type I error: linear detrending literally hides the step.

As a conclusion of these experiments, the use of only such basic patterns for change as trend and step cannot account for the change pattern when it is not a linear trend nor a step change, even in the simple case where it is merely the sum of the two. To describe real shift, a method that considers shifts of any length is thus needed. A first step in that direction is still to understand how they can be differentiated more efficiently than with the approach presented in section 2.3.1.

2.3.3 Using test statistics to distinguish trend from step

The question asked in this section is: can we distinguish linear trend from abrupt step by using the MK and Pettitt statistics, as announced in Figure 2.1? Let us now go further with looking at the MK statistics S not just as a way to only determine if there is a systematic change in a time-series, but also as carrying information about the timing and duration of this change. From the definition of S in equation (2.2), $\forall t$, $1 \leq k \leq n$ (where n is the lenght of the time-series), we have the following decompositions of S:

$$\begin{aligned} S &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) \\ &= \sum_{i=1}^{t} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) + \sum_{i=t+1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) \\ &= \sum_{i=1}^{t-1} \sum_{j=i+1}^{t} \operatorname{sgn}(x_j - x_i) + \sum_{i=1}^{t} \sum_{j=t+1}^{n} \operatorname{sgn}(x_j - x_i) + \sum_{i=t+1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) \end{aligned} \tag{2.11}$$

where the middle term can be recognized as the quantity introduced by equation (2.6) and taken at date t, while the two other terms look like MK statistics. This directly leads to the following relationship between the MK and Pettitt point statistics:

$$S = \sum_{1 \leq i < j \leq t} \operatorname{sgn}(x_j - x_i) + k(t) + \sum_{t+1 \leq i < j \leq n} \operatorname{sgn}(x_j - x_i) \tag{2.12}$$

Let us now introduce the Mann-Kendall statistics as computed over a subset of a time-series of size n. Then $\forall (p, q), 1 \leq p < q \leq n$:

$$S(p, q) = \sum_{p \leq i < j \leq q} \operatorname{sgn}(x_j - x_i) \tag{2.13}$$

so that equation (2.12) can be written as:

$$S(1, n) = S(1, t) + k(t) + S(t + 1, n) \tag{2.14}$$

A matricial interpretation of this result can help us understand what it means. Let us have the square matrix A of size $n \times n$ defined by:

$$a_{ij} = \begin{cases} \operatorname{sgn}(x_j - x_i) & \text{if} \quad j > i \\ 0 & \text{if} \quad j \leq i \end{cases} \tag{2.15}$$

A is an upper triangular matrix that also has zeroes on the diagonal. Using equations (2.2) and (2.15), the MK statistics is merely computed as the sum of the terms in that upper triangle. Its decomposition by equation (2.13) can be graphically translated by Figure 2.5, which helps to understand even better what the difference between a gradual and a step change may be. For a step, $S(1,t)$ and $S(t+1,n)$ are not supposed to carry any relevant information if $t = T$, because there is no change in the t first points nor in the $n-t$ last points. We can expect $k(T)/|S(1,n)|$ (which is noted as $K/|S|$ for simplicity) to be around 1. However for a trend, there is still change in the t first points and in the $n-t$ last points which carry some information. We can also expect $K/|S|$ to be smaller than 1. $S(1,t) + S(t+1,n)$ is greater in the case of a gradual change, so that from equation (2.14) we can expect $K/|S|$ to be greater in the case of a step than in the case of a gradual change.

Monte-Carlo simulations were carried out to test the reality of this theoretical difference, and the results are provided in Figure 2.6. On the left the slope is $\beta = 0.03$ for both trend and step, like in section 2.3.1, while it is $\beta = 0.06$ on the right of Figure 2.6. For the step, a greater portion of the information can be accounted for at a single point than for a trend. For $\beta = 0.03$, around 63% of the steps have $K/|S| > 1$ while around 88.5% of the linear trends have $K/|S| < 1$. For $\beta = 0.06$ the difference is even wider between trend and step. The distinction between trend and step cannot be perfect because of the noise, but using the MK and Pettitt statistics leads to better perceiving the difference.

Thus, using the statistics of the Pettitt and MK tests together can differentiate trends from steps much better than when deriving only the significance of change from these same test statistics. In the following this relationship is further extended to account for the duration of change as well.

2.4 An integrated framework for change assessment

This section, the core part of this chapter, presents the method that goes beyond the recognition of such basic patterns as step changes or linear trends, in order to deal with complex patterns in real data. As outlined in Figure 2.1, we are dealing with the detection of both timing and duration of change (section 2.4.1) but also with the issue of the detection of its statistical significance, which is intertwined with that of serial correlation (section 2.4.2). These aspects are integrated in section 2.4.3, and the issue of the detection of current change is tackled last (section 2.4.4).

2.4.1 Interpreting the Mann-Kendall statistics

If the shift of the data is more gradual than a single-point change, the rectangle in Figure 2.5 should be bigger for the sum of the terms it encompasses to be as big as $|S|$. For $0 \leq X \leq |S|$, the question is how big would it need to be so that $k > X$? In other words, how small the dotted triangles would need to be to contain less than $|S| - X$? As represented on Figure 2.7, these triangles would then represent $S(1,t)$ and $S(t+d+1,n)$, with the difference between t and $t+d$ being the period when change occurs (at the level X). As for d, it can then be interpreted as the duration of change. The information about change would then be encapsulated into the following Pettitt period statistics (also represented on Figure 2.7):

$$\begin{aligned} k(t,d) &= \sum_{i=1}^{t+d}\sum_{j=t+1}^{n} a_{ij} \\ &= \sum_{i=1}^{t}\sum_{j=t+d+1}^{n} \operatorname{sgn}(x_j - x_i) + \sum_{i=t+1}^{t+d-1}\sum_{j=i+1}^{t+d} \operatorname{sgn}(x_j - x_i) \\ &\quad + \sum_{i=1}^{t}\sum_{j=t+1}^{t+d} \operatorname{sgn}(x_j - x_i) + \sum_{i=t+1}^{t+d}\sum_{j=t+d+1}^{n} \operatorname{sgn}(x_j - x_i) \end{aligned} \tag{2.16}$$

so that the first term of the decomposition (designated by (i) on Figure 2.7) of $k(t,d)$ compares a period "before" change (from 1 to t) to a period "after" change (from $t+d+1$ to n), while the last two ones respectively compare the periods "before" (iii) and "after" (iv) to a period "during" change (from $t+1$ to n). Finally, referring to equation (2.13), the second term of (2.16) can be identified as $S(t+1,t+d)$ (ii). It carries information on the increase (resp. the decrease) during the relevant change period, during which the measurements are ordered in increasing (resp. decreasing) order for an increase (resp. decrease). For any duration d, we can use $k(t,d)$ in the same way as $k(t)$ defined by equation (2.6) to introduce a modified Pettitt statistics:

$$K(d) = \max_t(|k(t,d)|) \tag{2.17}$$

This statistics, unlike the one defined in equation (2.8), can help determine the duration of change besides its timing. For any d, the associated date of change t_c is thus:

$$t_c = \arg\max\{|k(t,d)|\} \tag{2.18}$$

so that $|k(t_c,d)| = K(d)$. Now the smallest duration for which the modified Pettitt point statistics can explain a fraction X of change is expressed as:

$$d_c(X) = \arg\min_d \{d\,|K(d) \geq X\} \tag{2.19}$$

The relevant timing and duration of change are obtained following the algorithm depicted in Figure 2.8. After picking X, the duration of change d is incremented until the corresponding $K(d)$ satisfies to equation (2.19). This ensures that the algorithm finds the smallest duration of change for which the Pettitt period statistics gathers the desired amount of information. Then it is possible to define the relevant period of change p_c putting together these notions of timing and duration from equations (2.18) and (2.19):

$$p_c = [t_c+1, t_c+d_c] \tag{2.20}$$

All the values between 1 and n are possible for p_c. We are no longer in a situation where a monotonic linear trend and an abrupt step are the only basic change patterns used to describe the data. Instead, a broad range of gradual changes can be detected between these two extremes. In this new configuration, the generalized Pettitt statistics is still related to the MK statistics by:

$$S = S(1, t_c) + k(t_c, d_c) + S(t_c + d_c + 1, n) \tag{2.21}$$

In particular, if $X = 1$ then from equations (2.21) and (2.19) we have $S(1, t_c) + S(t_c + d_c + 1, n) < 0$ if the values of the data are increasing (> 0 if these same values were decreasing). $k(t_c, d_c)$ can then be interpreted as carrying all the information about change. In other words, our interpretation of the MK statistics allows to give a timing and duration for the relevant change pattern in the data.

However, that doesn't mean there is no information left about change at the beginning or at the end of the series. For example, in a time-series featuring a significant increase, we can have $-S(1, t_c) > S(t_c + d_c + 1, n) > 0$ so that the decrease at the beginning of the record offsets the increase at its end. Then $d_c$may be underestimated. That is why, knowing the normal quantile of the MK statistics in the parts before and after change can be useful.

2.4.2 Impact of serial correlation on test significance

While the above section explained how relating the Pettitt and MK statistics can inform on the timing and duration of change, knowing whether that change is statistically significant remains an issue. The tests themselves have to be used. Most statistical testing methods were derived using the hypothesis of independent data, and the ones used in this thesis are no exception. Correlation creates redundancy between distinct observations, so that the useful sample size can be considerably less than the number of observations (*Koutsoyiannis and Montanari*, 2007). We saw in section 2.2 that for both the MK and Pettitt test, in a shorter time-series the observed variations are more likely to be due to chance (random events) alone. Thus when serial correlation is present, the confidence intervals associated with the different statistics are modified.

For a time-series of size n, an estimate of autocorrelation at lag k is given by (e.g. *Box and*

Jenkins, 1976):

$$r_k = \frac{\frac{1}{n-k}\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.22}$$

Autocorrelation is deemed significant at the level α if it falls outside the following interval:

$$\frac{-1 + z_{\alpha/2}\sqrt{n-2}}{n-2} \leq r_k \leq \frac{-1 + z_{1-\alpha/2}\sqrt{n-2}}{n-2} \tag{2.23}$$

In this thesis we will choose a 10% significance level for the autocorrelation estimates, so that $z_{1-\alpha/2} = 1.645$.

This section examines the case of serial correlation in the cases of the MK and Pettitt tests before detailing the two opposites interpretations that may be given when fluctuations are observed in a time-series.

A) Mann-Kendall test and serial correlation

In the presence of serial correlations, the variance of the MK statistics is modified. Indeed positive correlation increases $V(S)$ in equation (2.4) while negative correlation tends to decrease it (*Hamed and Rao*, 1998). In this section, the cases of first- and higher-order correlation are discussed separately.

(i) First order serial correlation: *Von Storch* (1995) acknowledges the importance of that issue for climatic time-series and proposes prewhitening to deal with correlation structures, only in the particular case of a stationary AR(1) process (for stationary autoregressive processes, see *Box and Jenkins*, 1976), described by the following relationship between two consecutive observations:

$$X_{i+1} = \rho_1 X_i + \epsilon_i \tag{2.24}$$

where $-1 < \rho_1 < 1$, and the (ϵ_i) are independent and identically distributed stochastic variables. As demonstrated by *Yue et al.* (2002b) and *Yue and Wang* (2002c), prewhitening effectively removes the AR(1) component. Nevertheless, these studies also show that when a trend is superimposed to an AR(1) process, prewhitening modifies the value of the slope of the trend. In addition, if the trend is removed first (using the same Sen-Theil estimate as in the present work) we can still provide a good estimate of the AR(1) component. Based on these three observations, they propose a procedure where the AR(1) component is removed after the trend present in the data. This procedure requires to compute the Sen-Theil estimate of the slope, β, as in equation (2.10). Then the slope is removed from the series (X_i) to give a detrended series (Y_i), with:

$$Y_i = X_i - (\beta \times i) \tag{2.25}$$

r_1 is calculated for the detrended series, which for $i > 1$ is then prewhitened as follows:

$$Z_i = Y_i - r_1 Y_{i-1} \tag{2.26}$$

and the MK test is finally performed on the series:

$$X_i' = Z_i + \beta \times i \tag{2.27}$$

In fact, *Yue et al.* (2002b) also note that in most hydrological time series featuring annual data, the serial correlation coefficients are relatively weak. As a consequence, because of equation (2.23), the autocorrelation coefficients $r_k = \rho_1^k$ are such that only the first one is significant. This means we can apply the procedure these authors develop if and only if r_1 is the only significant autocorrelation coefficient.

The only issue that has not been dealt with in the two aforementioned studies is whether applying this trend-free prewhitening procedure on series that don't have significant trends can alter the estimate of the AR(1) component. This is relevant question as the significance is assessed after trend removal. Monte-Carlo simulations were performed to check that trend removal on a pure AR(1) doesn't significantly affect correlation. Equation (2.24) is applied on standard normal processes, with values of ρ_1 being incremented from 0.15 to 0.9 by intervals of 0.05. The results are presented in Figure 2.9, using the mean values of autocorrelation coefficients estimated via equation (2.22) before and after trend removal. We can notice that a systematic bias was introduced in the measure of serial correlation, even before removing the trend. This is because while the mean of the process used to generate the data is 0, in general the estimated mean is not 0. This leads to underestimating serial correlation. At all correlation levels, this underestimation bias is greater than the supplementary bias introduced by trend removal. Further, the mean of relative error induced by detrending is always less than 10%, and is less than 5% for $\rho_1 \geq 0.3$. This suggests that we can use MK2 even when no actual trend is present.

From now on, we define the procedure set up by *Yue et al.* (2002b) and described by equations (2.25) to (2.27) as MK2, and use it when serial correlation is present only at lag 1.

(ii) Higher orders of serial correlation: Yet, in their study of streamflow trends in Indiana, *Kumar et al.* (2009) show that the aforementioned procedure is not effective when higher-order lags exist. They use another method that comes from a systematic study from *Hamed and Rao* (1998). These authors derive an alternative empirical formula for the calculation of variance using the equivalent sample size n^*:

$$V(S)^* = V(S)\frac{n}{n^*} \tag{2.28}$$

where n is the original sample size. The quotient from equation (2.28) can be approximated by an equation that takes into account all the significant autocorrelation coefficients:

$$\frac{n}{n^*} = 1 + \frac{2}{n(n-1)(n-2)}\sum_{i=1}^{n-1}(n-i)(n-i-1)(n-i-2)r_i \tag{2.29}$$

For convenience, from now we define the procedure described by equations (2.28) and (2.29) as MK3.

Before using this method however, we need to know whether we need to remove the actual change pattern before dealing with the correlation structure, like we do with MK2. Therefore, Monte-Carlo experiments are carried out with independent standard normal observations, and slopes ranging from 0.005 to 0.05 (by increments of 0.005) are superimposed. The original MK test results are compared to the MK3 method, as displayed in Figure 2.10. The data are independent and both methods should have the same result. However, the original MK test has a much greater power. This is especially true at the 99% confidence level, when the magnitude of change is three times the standard deviation of the series ($\beta = 0.03$) and over. Then the modified test has a power of 1 at the 95% confidence level, but under 0.1 at the 99% level. The fact of computing the correlation structure without detrending leads to confusing significant changes with serial correlation. Then change may not be detected even though it is present.

(iii) Addressing the impact of serial correlation on the MK test: Thus, when confronted to significant autocorrelation at lags greater than 1 MK3 should be used, but only after removing the relevant change pattern. It can be empirically derived using particular forms for the probability distribution of the residuals of the process, so that if the lag is only of order 1, we may rather use the MK2 test which has a theoretical basis.

B) Pettitt test and autocorrelation

Like for the MK test, the Pettitt test assumes independent data, and serial correlation may affect the computation of the test significance. Although search for a systematic study of the impact of serial correlation on the Pettit test in the literature proved unfruitful, it seemed relevant to test how it reacted to serial correlation, to see whether the test could be used instead of MK in the presence of serial correlation.

To test the effect of positive correlations, the rejection rate for AR(1) processes was computed for the same series as in Figure 2.9. It is presented in Figure 2.11.a and proves that positive correlations favor the detection of change in the mean. This means we cannot use the Pettitt test as an alternative method when positive correlation exists. To test the effect of negative correlations, the power of the test was computed for series with a centered step change of the same magnitude as the standard deviation ($\sigma = 1$), again using Monte-Carlo experiments. An AR(1) process with ρ_1 ranging from 0 to -0.9 (by steps of -0.05) is superimposed on the series. Results are given by Figure 2.11.b and prove that the power of the test is not affected by positive serial correlation if $\rho_1 \geq 0.4$. This suggests the Pettitt test may be used to detect change in the mean associated with residuals that exhibit negative correlation. For extremely negative correlations such as those represented in Figure 2.11.c, the power of the test decreases drastically because the negative AR(1) process imposes oscillations of greater magnitude than the step change itself. But annual hydrological series are unlikely to have such a behavior (*Yue et al.*, 2002b).

This means we can use the Pettitt test only with relatively weak negative correlations, but never with positive ones. Given the absence of any systematic method to avoid the impact on correlation involved in the Pettitt, we only use the MK test with its MK2 and MK3 variants when dealing with the significance of change.

C) Two ways to look at variations in a time-series

Deciding whether observed variations describe physically meaningful change or arise from serial correlation is all a matter of interpretation. Some may seek to take into account as much correlation as possible to make their tests more conservative (e.g *Douglas et al.*, 2000; *Kumar et al.*, 2009). Thus when change is detected, it then becomes more robust. But the goal of trend assessment is also to detect change to provide information to decision makers (*Yue et al.*, 2002b), so that missing some relevant change patterns by using very conservative methods can have its drawbacks. While setting up a methodology dealing with non-stationary assessment, we mean to avoid mistaking meaningful change for serial correlation. Thus we would advocate removal of the relevant change patterns prior to dealing with serial correlation.

2.4.3 Integrated procedure for assessing changes in streamflow

Let us now incorporate sections 2.4.1 and 2.4.2 into a procedure to address the timing, duration and significance of change in a time-series at the same time. It is outlined in Figure 2.12. From now on, it will be defined as PCPS (Procedure for Change Pattern and Significance).

Step 1 is to detect the relevant change pattern following the algorithm described in Figure 2.8. Any levels of X can be used, but it is recommended to choose $X = |S|$ as one of the levels, because ultimately we want to account for 100% of the MK statistics. Picking $X > |S|$ is not physically meaningful and should therefore be avoided.

The pattern detected in step 1 should be considered when removing the change pattern present in the data, prior to computing serial correlation. That is why trend removal comes as step 2, after getting p_c. There are many possible ways to do it. A simple one is to use a linear trend, for example via computing the Sen-Theil β to fit to each of the 3 sub-periods: before, during and after change. That is what is being done in Chapter 3.

Step 3 consists of assessing the correlation structure of the residuals. Its result leads to the choice of the appropriate version of the MK test in Step 4. This leads to the determination of the significance of the observed change pattern at the desired significance level, following the learnings from section 2.4.2 ($\alpha = 0.1$ in this thesis). Like in *Box and Jenkins* (1976) one can choose to discard the significant correlation lags that occur several points after the last significant autocorrelation coefficient, on the basis that they don't really represent serial dependence any more. In practice while applying PCPS in Chapter 3 all the coefficients after the second non-significant one are thus discarded.

Then step 7, finding a suitable physical interpretation is essential: a statistical model is only valid if physically backed (e.g. *Koutsoyiannis and Montanari*, 2007). In order to do that, it may also be useful to have an idea whether current change may be a valid model or not (step 6). This is developed further now.

2.4.4 The possibility of current change

One of the most useful pieces of information about change is to know whether it is still currently going on. Indeed ongoing change can be expected to continue in the future, with all the potential implication this entails. However, statistical analysis alone cannot tell whether change is going on: for that we have to find the process at its origin. We can only have tools that may provide hints as to know whether current change is a reasonable model or not. Two such conceptual tools are described as follows.

A) The concept of strong trend

The first one aims at improving the method proposed by *Villarini et al.* (2009a), which is already discussed in section 2.3.1. It supposes that current change is a reasonable model if, after the detection of statistically significant change point, a trend test applied to the period starting after that change-point can detect change at the desired significance level. Using more complete description of the change pattern provided by PCPS, it is proposed here that a trend analysis should be performed excluding only the period “before” change (from 1 to t_c). The shaded area of Figure 2.7 is then extended to include $S(t+d+1, n)$. We are left with only a “before change” period excluded for trend analysis, and a “during change” period, separated by t_c. The advantage of doing so is illustrated by Figure 2.13, representing the annual daily maximum at the Poplar Creek streamflow gauge at Elgin (for precisions on this USGS-operated gauge, please refer to Table 3.2 and Figure 3.2). The Pettitt test returns the year 1972, and no trend is statistically significant ($\alpha = 0.05$) for the 1973-2008 period. On the contrary, taking into account that changes starts on 1960 leads to detect significant change on 1960-2008, which fits much better both the data and the context of progressive urbanization of this watershed. If the trends are still significant at the level $\alpha = 0.05$ when taking the low (resp. high) values out the first part of an increasing (resp. decreasing) series, we call them “strong trends”.

One could object that the proposed method could lead to detect current change more easily for series in which change is historical. For instance, it would detect current change for the annual daily minimum at the Salt Creek gauge on Figure 2.14 (again, see Table 3.2), while it looks obvious that the time-series displays in fact a past increase. That is why one should also check whether the residuals detrended in that way are significantly more serially correlated than when considering the three periods before, during and after change. If they are, this would suggest an ongoing shift is not as good a model as a past one. If they are not, then we also want to find whether change is still significant if we exclude the “before” period. The rationale behind this is that if a linear trend excluding the period before change is a model that both displays significant change and embraces the data well enough that its residuals are not excessively correlated, then it is reasonable. The method relying on these two conditions would reject the hypothesis of current change in the case of Salt Creek.

The efficiency of the proposed method is compared to that of the one proposed by *Villarini et al.* (2009a) in Figure 2.18 for the same series as those used in Figure 2.6. The proposed improvement performs much better at telling whether a linear trend is an acceptable model.

B) Ending time of change

This method relies on checking which is the smallest t for which $S(t,n)$ and S have opposite signs. If the time-series documents an increase of the studied quantity, we then search an ending time t_e such that:

$$t_e = \min\{t|S(t,n) \leq 0\} \tag{2.30}$$

where "$\leq$" should be replaced with "$\geq$" in case of a decrease. Then, because the MK statistics is centered on 0, that can be interpreted as the $n-t$ last points only showing an internal change pattern opposite to that of the whole series. The possibility for change to be ongoing would then depend on how big $n-t$ is. Symmetrically, we can use this method to find out how many points at the beginning of the time-series don't show the same change patterns as the whole series, and define $t_b = \max\{t|S(1,t) \leq 0\}$ for an increase (< 0 for a decrease). Again for the same series as in Figure 2.6, the results for this method are presented in Figure 2.19. This illustrates the difficulty of the task of determining whether change begins before the time-series ends. For instance, having $S(1,t) < 0$ with $t \geq 20$ in the case of a linear trend of $\beta = 0.03$ is only possible in about 60% of the cases.

2.5 Applicability and limitations

After setting up the PCPS and explaining its theoretical basis, its strengths and weaknesses have to be explored, to determine exactly how useful it can be in a statistical exploration of hydrologic non-stationarity in a given area. By construction it is exact in the absence of noise at the level $X = |S|$. This has been verified by Monte-Carlo simulation on predetermined patterns. One can also check on real data with low variability that at the level $X = |S|$, the change patterns found by PCPS do reflect the timing and duration of change. Thus, Figure 2.15 shows how a step change is detected in streamflow can be detected on the Kishwaukee River, an agricultural watershed West of Chicago. Likewise, Figures 2.16 and 2.17 show how PCPS accurately detects gradual changes in the 7-day minimum flow of two Chicagoan watersheds, Des Plaines River and Salt Creek (the streamflow data used in this thesis is described in Figure 3.2 and Table 3.4).

From its theoretical basis, PCPS does not aim at returning the exact pattern in presence of noise. Two reasons can explain that. First, methods that differentiate signal from noise in a short climatic time-series already exist (*Ghil et al.*, 2002), for instance singular-spectrum analysis (*Vautard et al.*, 1992). Second, treating variations in hydrologic time-series as noise is only an interpretation which comes from electrical engineering but may not be meaningful in geophysical sciences (*Koutsoyiannis and Montanari*, 2007). Other interpretations, such as long-range memory, have been suggested since the pioneering work by *Mandelbrot and Wallis* (1968). PCPS aims at providing a description of the change pattern that takes into account the contributions from different sources such as climate or urbanization, without making any assumption about the stochastic processes that may generate those contributions.

Thus, as far the author of this thesis is concerned, the two conditions that validate PCPS should be that 1) it returns an exact result on predetermined patterns and 2) it proves to be useful for the spatial and temporal analysis of change. This section is to further understand how stochastic perturbations interfere with pattern detection.

Because it is a rank-based method, it tends to return the length of the longest change pattern when several are present in the data, and not necessarily the length of the most obvious one.

Figure 2.20 illustrates that. This poses the problem of the behavior of the method when confronted to stochastic perturbations (such as noise). It can make the detection of the duration of change more difficult, but also impact more shifts of smaller magnitudes. Understanding the impact of noise is essential because the PCPS is most useful for noisy-time series. On the contrary, if the change pattern is obvious, *Kundzewicz and Robson* (2004) argue that the use of any sophisticated methodology is useless since looking at the time-series is enough.

Three questions are going to be addressed here. 1) How sensitive is this method to different change patterns? 2) How robust is it to noise and to low-intensity processes? 3) How much can this method be relied on to detect the different processes involved in a spatial and temporal analysis of streamflow change? Because identifying the true duration of change (question 1) deals with sensitivity, which is contrary to robustness (question2), these two topics are going to be tackled together. Question 3 will then be dealt with separately.

2.5.1 Noise robustness and pattern sensitivity

Figures 2.6, 2.18 or 2.19 already suggested that there is not always that much difference between a noisy linear trend and a noisy shift, even with 100 data points. For example, Figure 2.19 shows that in the case of a linear trend, there is a variable number $n - t_e$ of points over which no upward change is detectable. Similarly, there is also a variable number t_b of points at the beginning of the series for which no trend is detectable. This means that for the linear trends of Figure 2.19, the period of change p_c that can be found is smaller than $\max(t_e - t_b, 0)$ which is variable quantity, even when the change period is the whole time-series (in the case of a linear trend).

Conversely, this means that the duration of change returned by the PCPS cannot perfectly represent the amount of time over which the mean of the underlying process shifts. As such, study of a single time-series cannot yield a robust result. This is not due to the method but the presence of noise in the time-series itself. Yet, one can test how well the method can approximate the "true" duration of change in the interval over which a shift is detectable. For the same time-series as Figure 2.19, the results are presented in Table 2.1. They suggest that the whole interval between t_b and t_e is not used by PCPS as the relevant time range over which the shift happens. This may be a drawback when the real change pattern is a linear trend, but in the case of a step change it allows the derivation of a better estimate than the raw $t_e - t_b$. Thus the PCPS achieves a trade-off between the detection of sudden and more gradual shifts.

But while this analysis only concerns step changes and linear trends so far, the proposed methodology is to deal with different and varied patterns. Let us look at how well the PCPS responds to different durations of change, by superposing a white noise ($\sigma = 2$) over processes that show a shift over different periods of time (left on Figure 2.21). With the chosen level of noise, differences between two of these change patterns are very difficult to detect by visual examination. Results from Monte-Carlo simulations show that the PCPS makes a difference, albeit once again without being able to uncover the "true" duration of change. Different levels of X are chosen. As expected from section 2.4.1, the smaller the amount of information X that we choose to gather, the smaller the duration of change the algorithm returns. The level $X = |S|$ is the one that is the most sensitive to the different change periods. It only features differences that are much smaller than the "true" ones. This also suggests that lengthy durations of change and large differences between time-series with similar noise (e.g., with similar climatic inputs) can be a sign of the existence of strong gradual change patterns. As discussed from Figure 2.19, these cannot be detected in a robust way from a single time-series, but the robustness of such signals can be enlightened by the examinations of many different time-series.

Before discussing the applicability of the PCPS to spatial analysis, we should check how the superposition of different patterns with distinct magnitudes is handled by this framework in presence of noise. Given that the detection of the most obvious ones is important for the robustness of the insights it can bring, low magnitude change can be undistinguishable from noise, and can sometimes arise from it. Let us pick up an idealized example derived from Chapter 3, where a climatic step change occurs alongside with gradual effects of urban development. Thus a white noise ($\sigma = 2$) is superposed over a time-series of $n = 60$ points in three scenarios:

- A) a step increase of magnitude 3 at $t = 20$;
- B) the same step change plus a linear increasing trend over the first 40 points;

- C) the same step change plus a linear increasing trend over the last 40 points.

The magnitude of the shift caused by the trend varies from 1 to 3 throughout the Monte-Carlo experiments. This aims at determining whether the detection of a long period of change is likely to arise from low intensity variations. The results are displayed on Figure 2.22. In all cases, the step is marked by a period of change starting before $t = 20$. As expected, the ending date is later in Scenario C, but the difference between B and C is minimal (3 points in average) when the magnitude of the trend is 1. Once again, the results suggest that noise reduces the differences between processes that display a different mean, so that very contrasted results from the PCPS could be the sign of the existence of robustly different processes.

2.5.2 Applicability to spatio-temporal analysis of change

As we can see from the previous section, the robustness of a signal detected on a time-series cannot be inferred from the examination of a single time-series. That is why PCPS is an exploratory tool to be used over datasets of numerous time-series, so as to provide information for a better understanding on the causes for hydrologic non-stationarity. Data exploration is an essential part of statistical analysis (*Kundzewicz and Robson*, 2004), which provides insights that can be useful to model development (e.g. *Kumar et al.*, 2009; *Villarini et al.*, 2009b).

Discovering temporal patterns at different locations can thus lead to useful insights into the spatial repartition of different causes for change. As previous experiments just suggest, a pattern is more robust if it is discovered across a whole set of time-series. The same is true for contrasted differences between two sets of time-series, which can for instance be linked to differences in the physical characteristics or in the history of these two sets. Because noise makes the distinction between two distinct patterns more difficult, large differences in temporal patterns are likely to be linked to distinct physical mechanisms at the spatial locations that can be associated to the studied time-series. Then, the quality of the results that PCPS gives ultimately lies in the meaningfulness of the insights it provides. Indeed, it should be noted that while hydrological series are stochastic processes, they also follow equations coming from physical principles. The possibility to model them as stochastic processes comes from the fact that those equations may be non-linear and chaotic (*Sivakumar*, 2008), but we must keep in mind that deterministic models can also (imperfectly) describe the behavior of streamflow series. This dual vision of hydrologic processes is advocated by *Vogel* (1999), and suggests that the variations and noise found in a hydrologic time-series can eventually make sense when traced back to their ultimate causes. The method presented in this paper is a tool to account for these variations and try to understand them. One can hope that variations that appear in many time-series might not merely represent noise, but can be ultimately linked to a physical explanation.

2.6 Proposed validation of the PCPS

The experiments described in section 2.5 are descriptive of the possibilities of PCPS, but they do not represent a rigorous validation of the method. This section is to briefly present an idea to achieve this latter objective.

As explained in section 2.5, different levels of X correspond to different ways to solve the trade-off between detecting step changes and gradual changes. Indeed a lower X more accurately detects step changes, but also tends to shorten the duration of gradual changes more than at the level $X = |S|$. In fact the level $X = 0$ corresponds to the Pettitt test as the algorithm described in Figure 2.8 stops at the first iteration. At $X = 0$ there is no search for the duration of change.

Thus, one could expect the level $X = |S|$ to contain all the change patterns which magnitude is high enough for them not to be hidden by noise. Justifying it by Monte-Carlo simulations would allow to ensure that in practice, PCPS does capture all the significant features of change. These simulations would consist in retrieving the dates of change for processes that are the sum of a standard normal white noise ϵ and an underlying change pattern c:

$$f(t) = c(t) + q\epsilon(t) \tag{2.31}$$

where q is incremented from 0 to 1.2 by steps of 0.1. This would then allow a better validation of how PCPS captures the relevant changes with different levels of noise. The considered patterns are:

- One and then several step changes, to test how much PCPS can detect them all.
- Gradual change of different lengths, on time-series of different lengths.
- Random combinations of step changes and gradual changes.

2.7 Summary and conclusions

This chapter studies two rank-based tests in a consistent framework: the MK trend test and Pettitt change-point test. They are usually merely used to detect significant change in time-series, but the non-stationarity issue in hydrology community requires to use these methods more efficiently. Experiments are conducted to better understand them and dissipate some common confusions: a trend test detects at systematic change over the whole time-series, but the change it detects does not have to occur on the whole length of the series; likewise, a change-point test detects if change occurs between the realizations of a process before and after a point, but it does not mean change occurred exclusively around this point. We show not only why they can't be used to remove change patterns separately, but how the statistics based on the method could be related and articulated together to provide an idea of the relevant pattern of change. This can be integrated into a rigorous and holistic methodological framework that incorporates the state of the art developments in dealing with the important issue of serial correlation. It uses the assessment of the timing and duration of change to detrend the data before looking for autocorrelation, thus avoiding to confuse it with meaningful change. Monte-Carlo simulations prov that the PCPS method, like any other, is limited in its detection capability by the presence of noise. As a consequence, though it is able to make a difference between different change patterns over a large number of simulations, this difference might be underestimated. This makes the assessments PCPS makes suitable for a spatial analysis of hydrologic non-stationarity: large differences in the timing and duration of change across several time-series are a robust indication of the existence of different underlying physical mechanisms for change. Chapter 3 will confirm this statement.

2.8 Tables and Figures

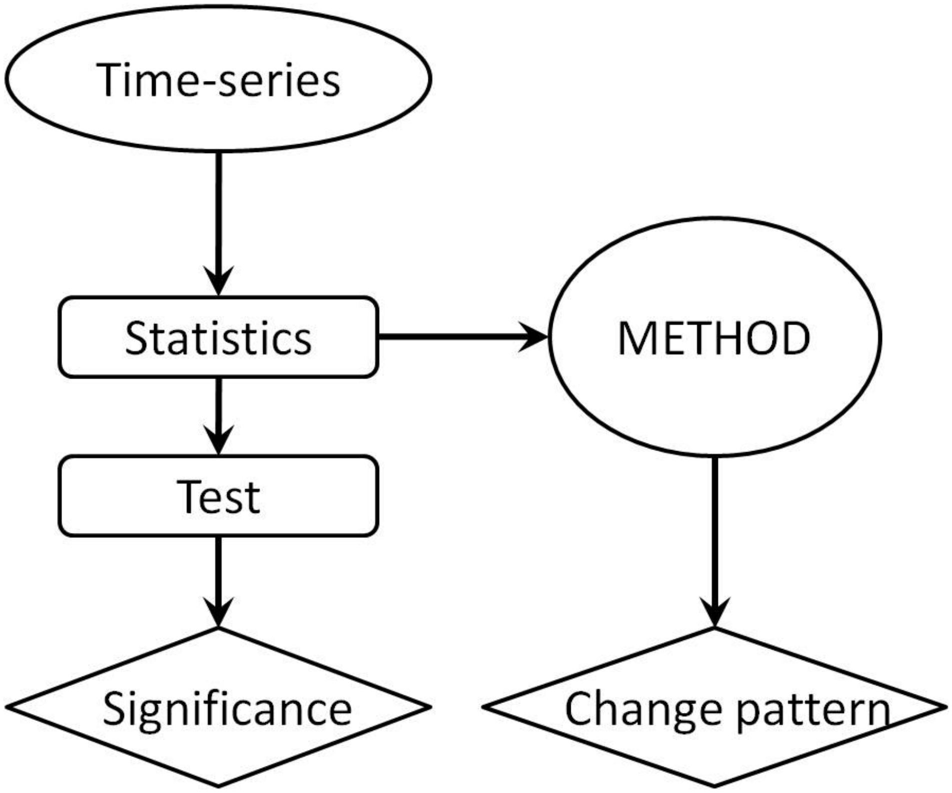

Figure 2.1: Conceptual framework for the methodology development. Like in previous works, significance is assessed through a statistical test, but its statistics will be directly used by a methodological development giving timing and significance of change.

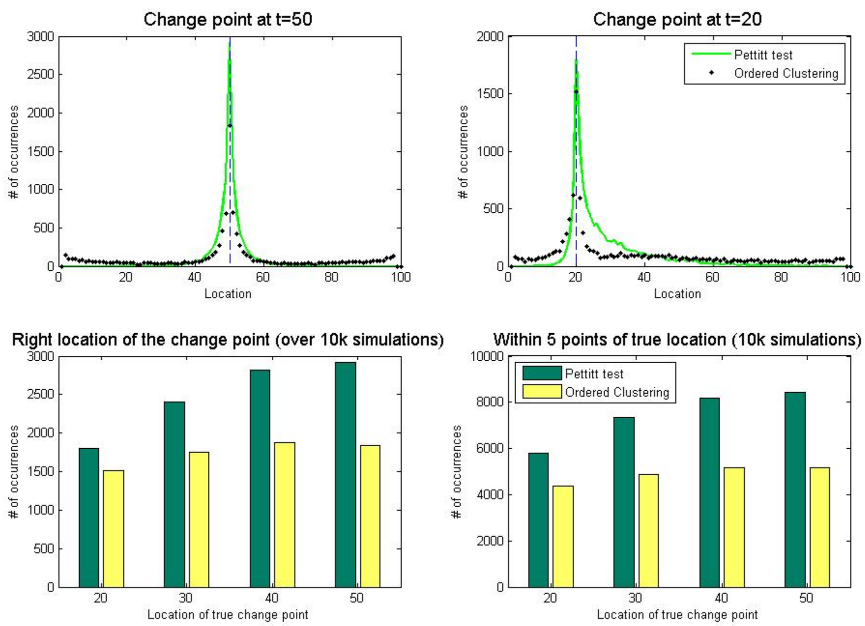

Figure 2.2: Comparison of the ability of the Pettitt test and an ordered clustering method to locate a step change. Ten thousand time-series with 100 points are simulated. The standard normal observations have a standard deviation of the magnitude of the step change. Step changes are successively set at $t = 20$, 30, 40 and 50.

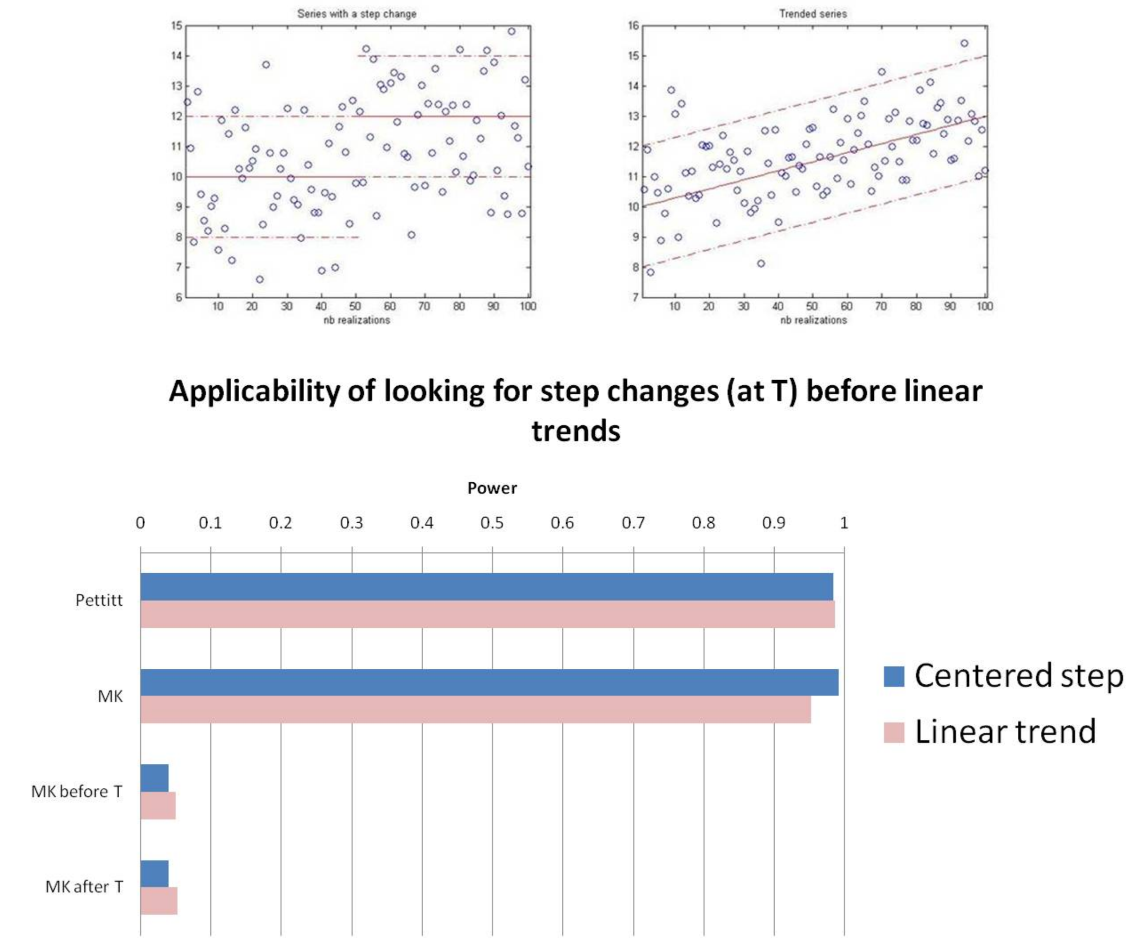

Figure 2.3: Top left (resp. right): series with a centered shift (linear trend). The red line represents the mean of the process while the dotted lines picture the standard deviation. Both top figures feature processes of same slope ($\beta = 0.03$). Bottom: performance of the Pettitt and MK tests with Monte-Carlo simulations with $10,000$ runs for both types of series shown above.

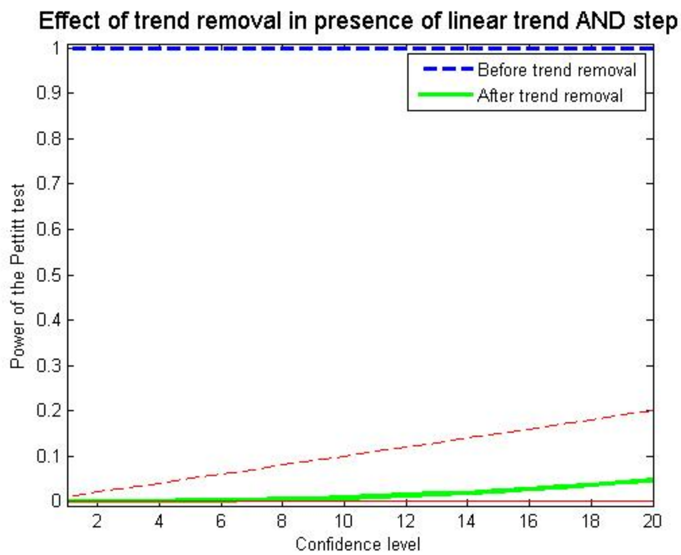

Figure 2.4: Comparison of the power of the Pettitt test before and after detrending a time series when both trend and step are present. In the detrended series the power of the Pettitt test is sensibly less than the type I error (dotted red line).

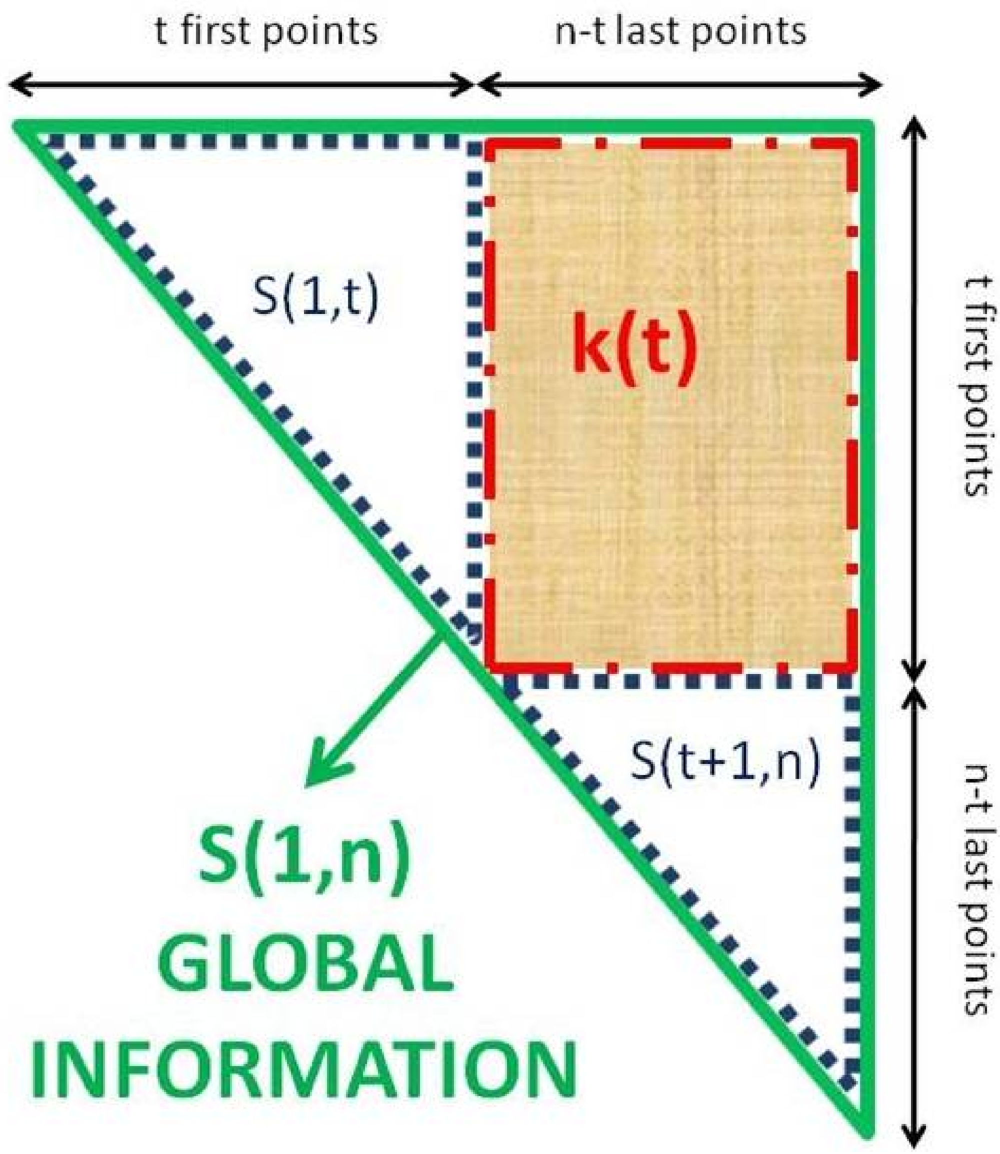

Figure 2.5: Matrix representation of the MK statistics S. While all the pair comparisons are within the green triangle, the red rectangle features those of the pair comparisons taken into account by the Pettitt statistics. The remnant is made of two upper parts of square matrices: the MK statistics at the beginning and end of the time-series.

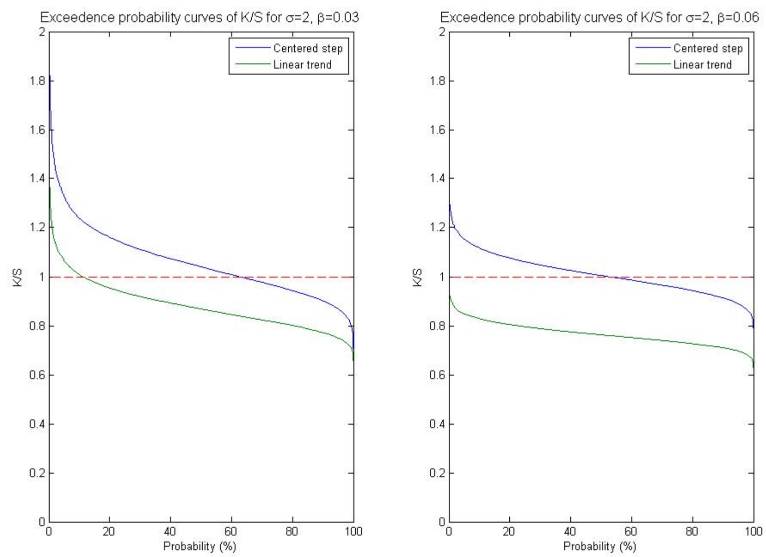

Figure 2.6: Measures of K/S for trends and shifts of the same linear slope. By representation with exceedence probability curves, we can see that for a step change, a single point explains more of the change.

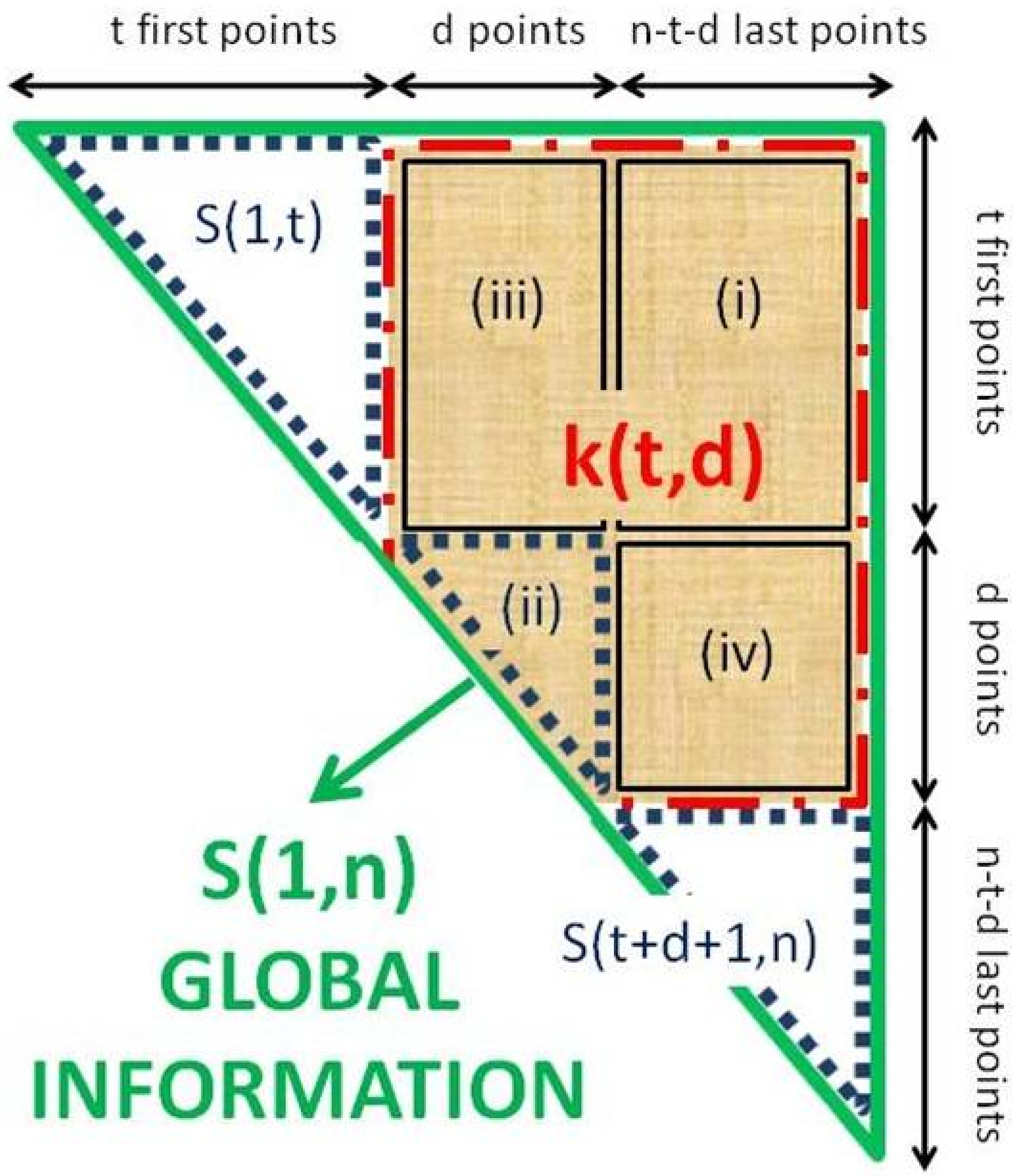

Figure 2.7: Matrix representation of S and of the Pettitt period statistics, $k(t, d)$. The decomposition follows the terms of the sum of equation (2.16).

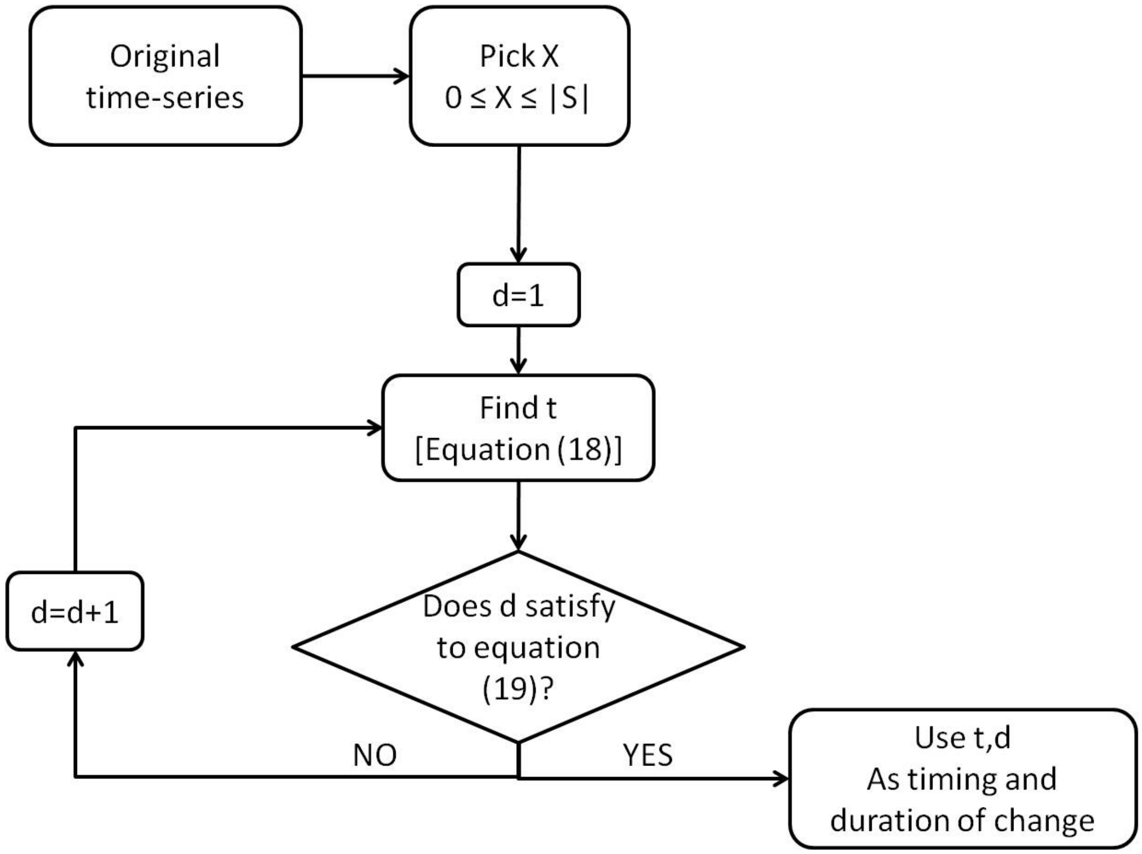

Figure 2.8: Proposed algorithm to identify timing and duration of change. The duration is incremented till $k(t_c, d)$ explains a fraction X of the information contained in S.

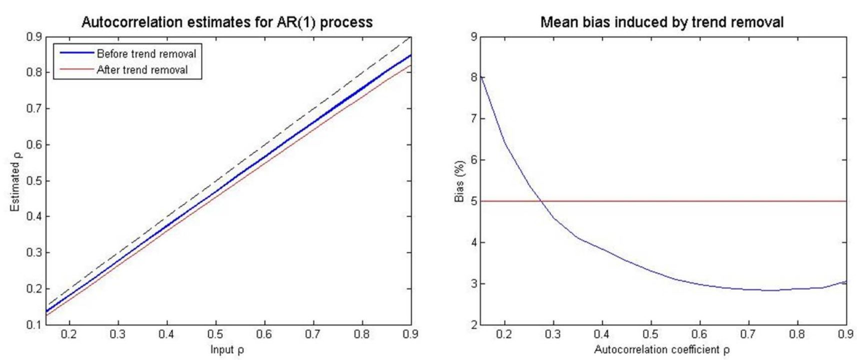

Figure 2.9: On the left: bias introduced by trend removal on the estimate of the first order autocorrelation coefficient of an AR(1) process. The dotted line is the $y = x$ line. Right: translation of the bias in terms of relative error.

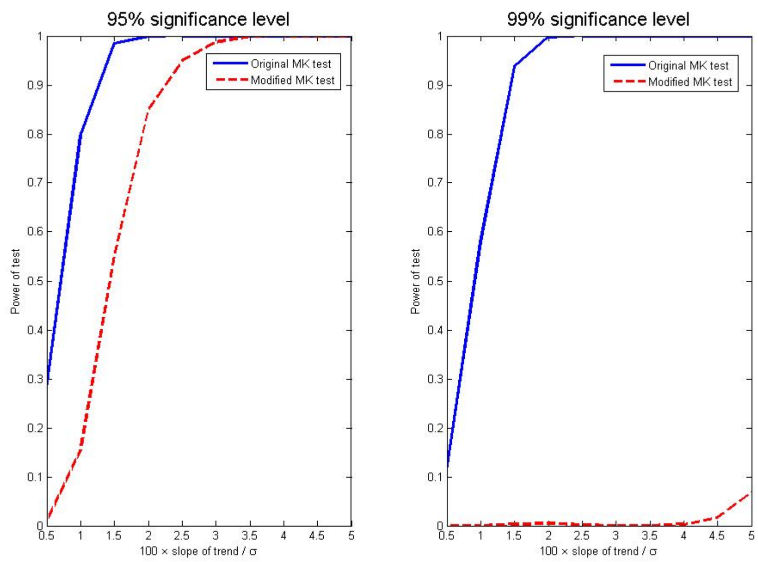

Figure 2.10: Compared power of the original MK test and the modified one (which accounts for serial correlation). A trend is superimposed on 100 independent standard normal realizations. This shows the interest of detrending first.

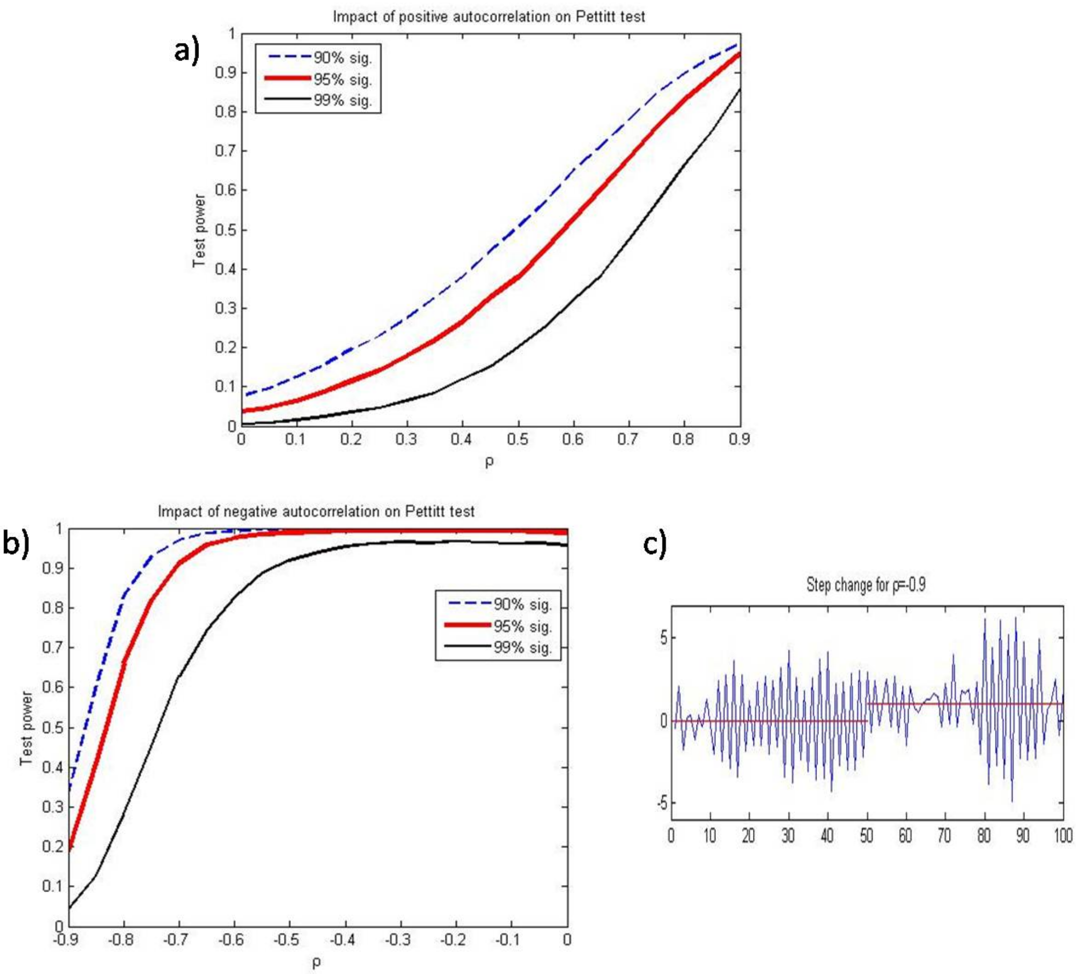

Figure 2.11: Power of the Pettitt test on AR(1) stationary processes for 100 data points, (a) shows positive serial correlation while (b) shows negative serial correlation. (c) shows an example of a time-series for $\rho = -0.9$.

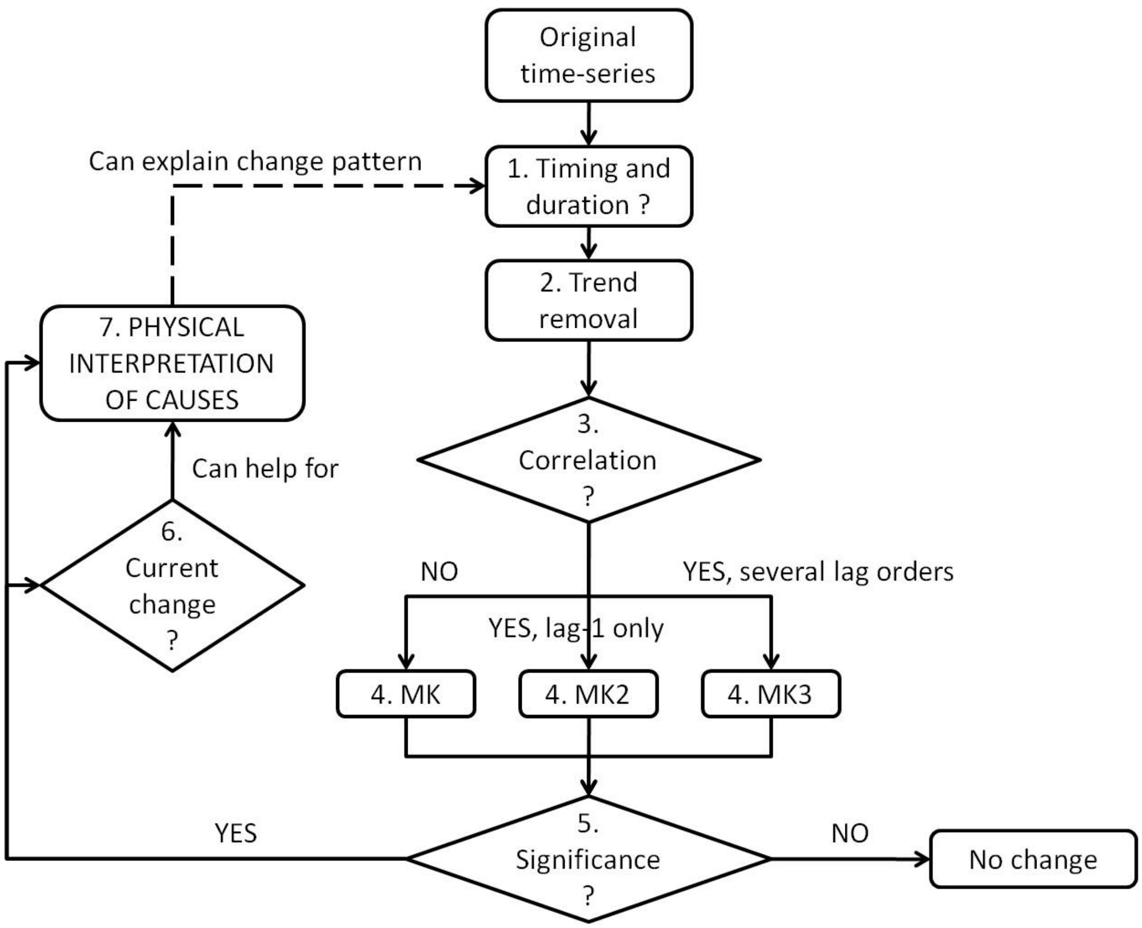

Figure 2.12: Procedure used to address timing, duration and significance of change. The change patterns discovered early are validated by the physical interpretation.

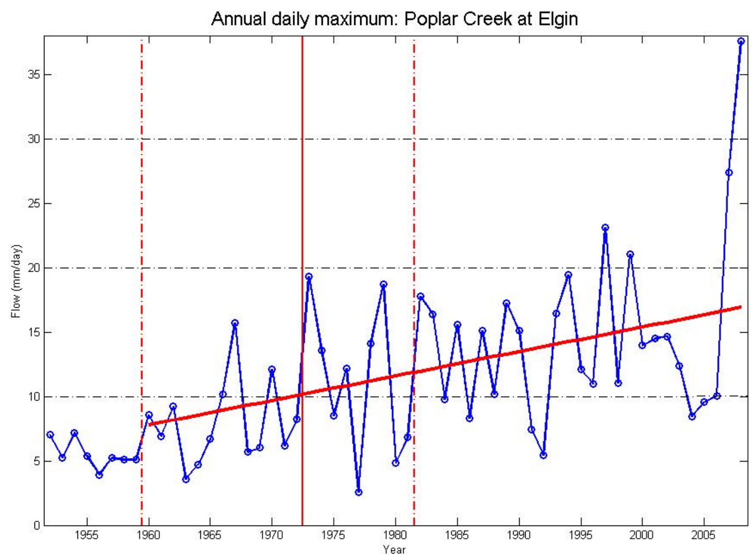

Figure 2.13: A change duration of 22 years is identified for the series for which it seems that urbanization is still driving an increase in peak discharge. The Pettitt year is the solid red line while the period of change is materialized by the dotted lines.

	β	$t_e - t_b$	d_c (X=1)
Step	0.03	14.2	8.4
	0.06	17.8	9.1
Linear trend	0.03	46.7	36.6
	0.06	68.1	61.0

Table 2.1: Duration of change (at the level $X = 1$) compared to the period upon which it is possible to detect an increase. The series are the same as for Figure 2.19 with 100 points long.

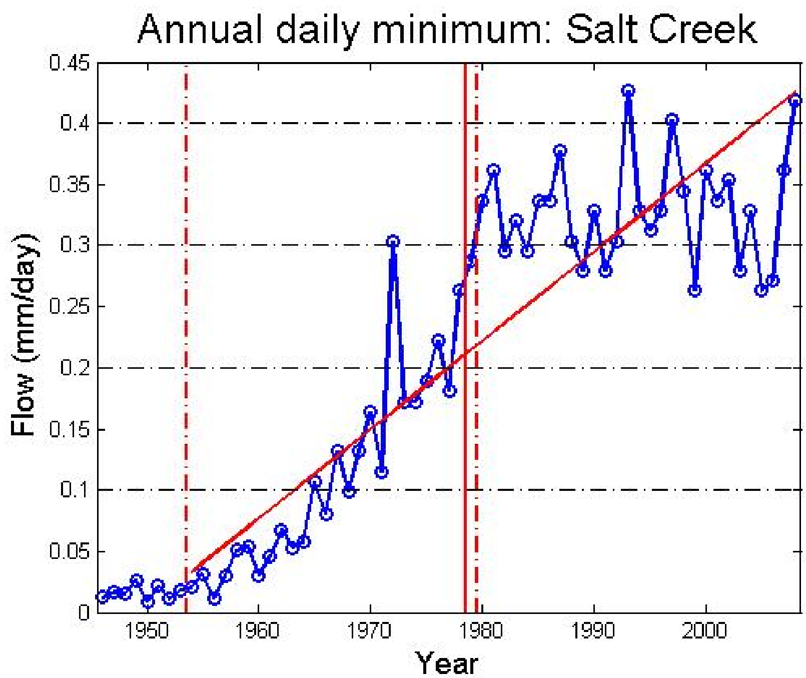

Figure 2.14: Example of a low flow series which describes an important historical change. A linear trend does not fit the data because change slows down and even stops.

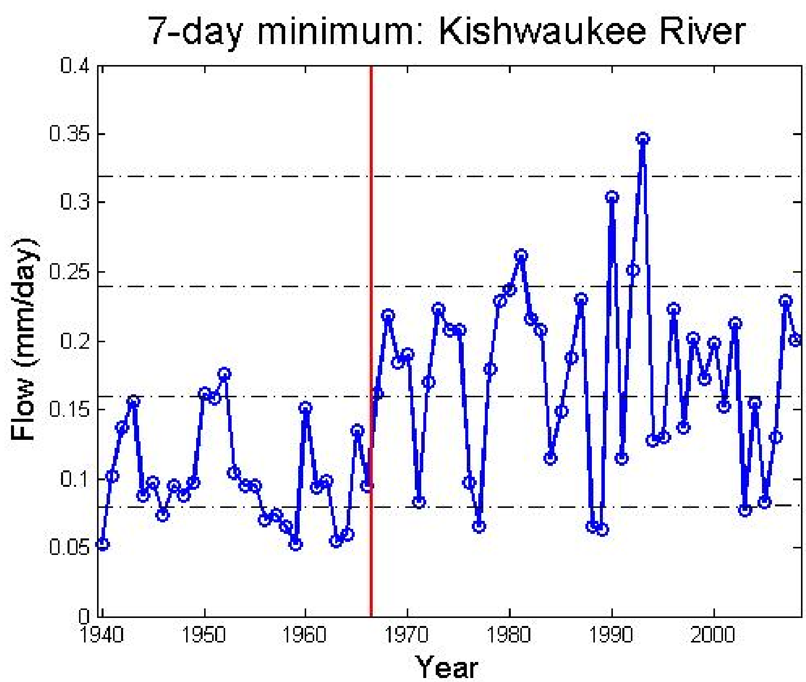

Figure 2.15: Dates of change in 7-day minimum: Kishwaukee River. This time-series features a step change in 1967 (vertical line).

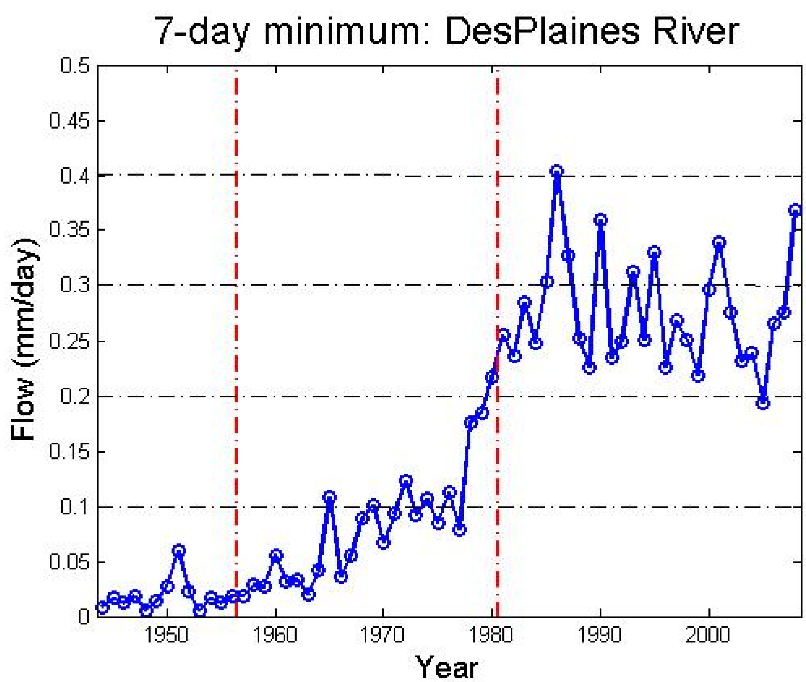

Figure 2.16: Dates of change in 7-day minimum: Des Plaines River. This time-series features a gradual change from 1957 to 1981 (dotted vertical lines).

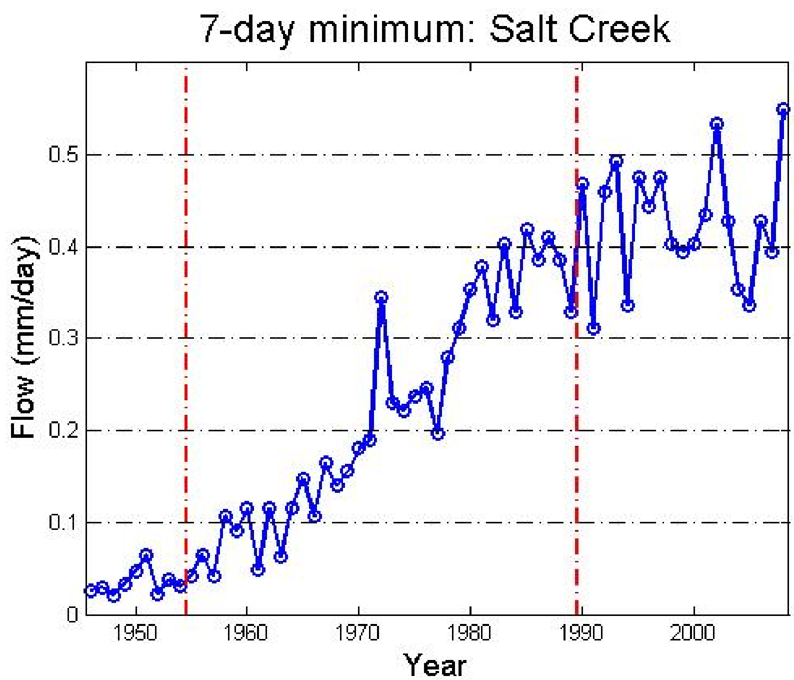

Figure 2.17: Dates of change in 7-day minimum: Salt Creek. This time-series features a gradual change from 1955 to 1990 (dotted vertical lines).

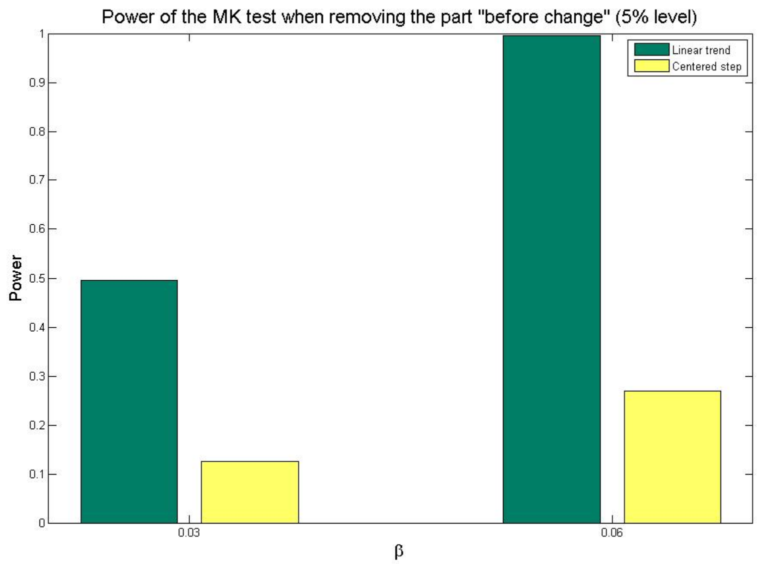

Figure 2.18: The figure shows how much trend and step can be differenciated by trend analysis even if the slope is weak, by removing the period “before change” found. Contrary to Figure 2.3 we have a small (at $\beta = 0.03$) but real difference.

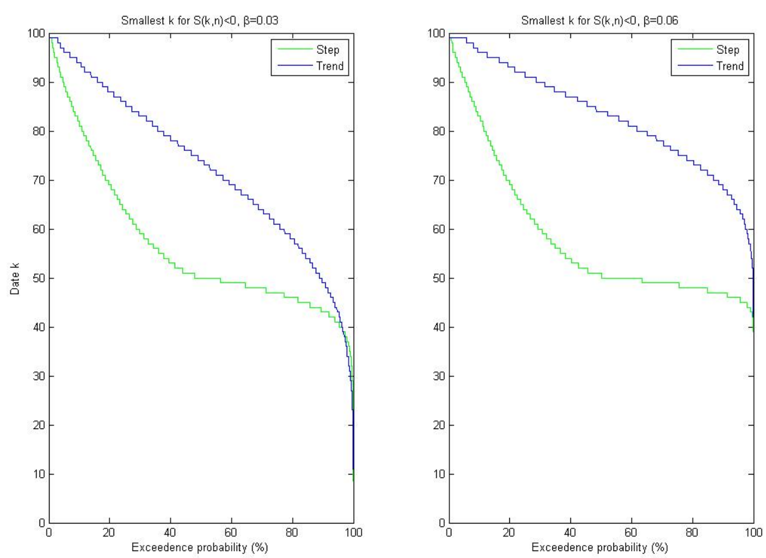

Figure 2.19: First point for which $S < 0$ while the series increases, for $\beta = 0.03$ and $\beta = 0.06$. This shows the difficulty of diagnosing current change from statistical analysis alone.

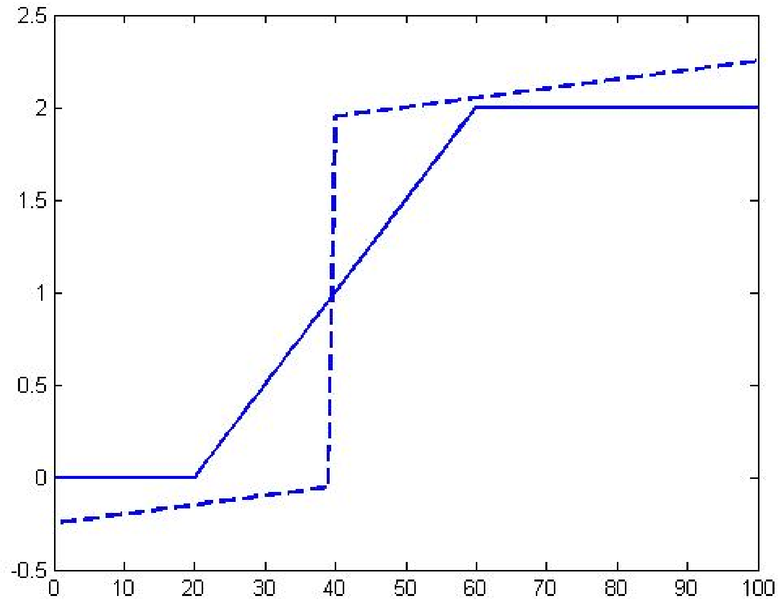

Figure 2.20: The PCPS detects change between point 21 and point 60 for the continuous line, and between point 1 and point 100 for the dotted line.

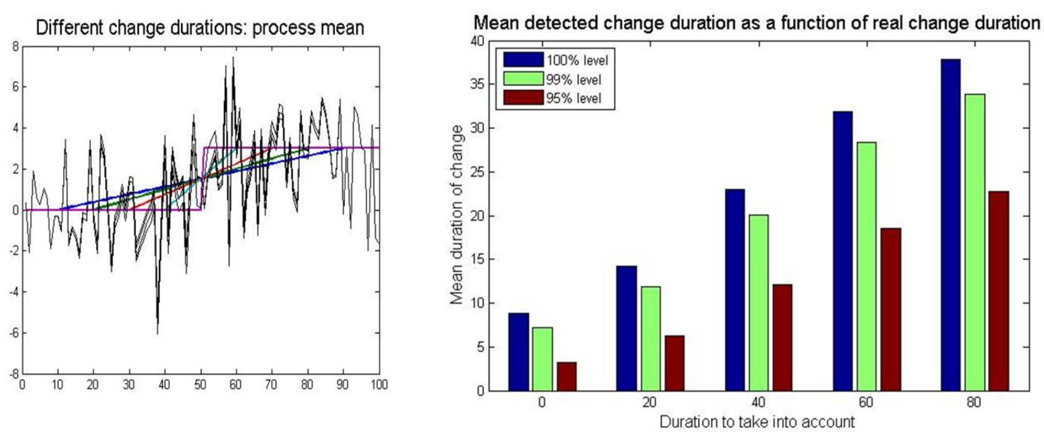

Figure 2.21: Right: mean change period found for change patterns of duration comprised between 0 and 80 points. They are centered and of same magnitude (left). In fact with a process standard deviation of 2, these processes are difficult to distinguish.

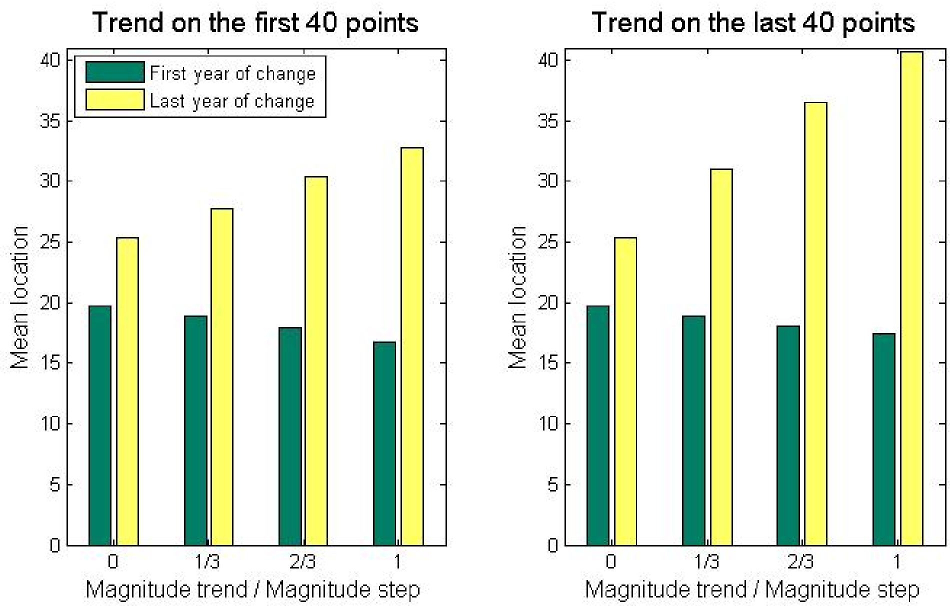

Figure 2.22: Scenario B on the left, Scenario C on the right. Both are compared to Scenario A (no trend). If the trend has a low magnitude, it is difficult to distinguish it from Scenario A.

Chapter 3

Impacts of urbanization and climatic variations in Chicago

3.1 Introduction

Now that we have established the PCPS method to help us extract meaningful and complex change patterns that could be present in the data, we are going to use it on the area we chose to focus on: urbanizing Northeastern Illinois. As stated in Chapter 1, the goal is to understand which effects of urbanization can play a foremost role in streamflow change at different scales within the Greater Chicago area.

The effects to consider in this Chicagoan context are potentially numerous and should be recalled before starting this analysis. We must differentiate the effects of land-use change from other effects of urban development. The basic effects of land-use change on the hydrologic cycle, through turning natural soils into impervious surfaces, have long been documented. Such surfaces let the water turn into runoff instead of infiltrating into the ground, thus enhancing flooding (*Hollis*, 1975; *Lazaro*, 1976) and reducing direct groundwater recharge, sometimes leading to baseflow depletion (*Ferguson and Suckling*, 1990). Suppression of vegetation can also lead to reduced evapotranspiration, which translates into an increased annual runoff (*Dow and DeWalle*, 2000). It can also translate into higher temperatures, which can locally have opposite effect on evaporation, especially during dry years (*Ferguson and Suckling*, 1990).

Other effects of urban development, distinct to those of land-use change, include stream channel alteration, interbasin transfers, effluent discharge, pumping. They can have opposite effects on baseflow and low flows (*Barringer et al.*, 1994; *Meyer*, 2005; *Claessens et al.*, 2006), while increases in peak flow can be locally mitigated by the design of detention basins (*Yeh and Labadie*, 1997), or by neighboring wetlands (*Burns et al.*, 2005). Other characteristics of the flow, like baseflow recession, can be affected by pumping and effluent discharge (*Wang and Cai*, 2009), or by detention basins (*Solo-Gabriele and Perkins*, 1997).

This study also considers the impact of climate fluctuations, a complicating factor when studying human interferences on the alteration of hydrologic processes such as streamflow (*Claessens et al.*, 2006; *McCormick et al.*, 2009). Rainfall increases, particularly those affecting extreme events, are reported in the Upper Midwest (*Karl and Knight*, 1998; *Pryor et al.*, 2009) and in Chicago (*Markus et al.*, 2007). A wetter climate has been reported to affect the low and mean flows of the Eastern United States in general (*Lins and Slack*, 1999; *McCabe and Wolock*, 2002) and of Illinois in particular (e.g. *Smith and Richman*, 1993). Finally seasonal changes can be detected at a regional level, especially for spring and fall (*Lettenmaier et al.*, 1994; *Groisman et al.*, 2001; *Small et al.*, 2006).

This chapter is as follows. In section 3.2, we present the study area and the data, and explain how the PCPS method was used. In section 3.3 the main patterns discovered in streamflow are presented. They will be compared with rainfall change patterns in section 3.4 so as to better understand what features of flow evolution can be directly attributed to climatic variability. Local and seasonal patterns will both be discussed. We can then try and relate to human interferences any unexplained change patterns observed in urban areas. We first deal with the evolution of maximum annual flow at a wide range of urban scales in section 3.5. Then in section 3.6, due to the complexity of the possible human interferences on low and mean flows, a city-wide water balance will be put in relation with the PCPS results. Finally in section 3.7, we will put those findings into perspective in the conclusion.

3.2 Study area and data

3.2.1 Study area

The study area mainly comprises the whole Chicago metropolitan area, as well as a few mainly rural basins surrounding it west and south. Lake Michigan is situated East and North of the city.

Regional hydroclimatology is reasonably wet and characterized by year-long rainfall (*Milly*, 1994a; *Sankarasubramanian and Vogel*, 2003) and an annual cycle in both precipitation and evaporation (*Milly*, 1994b). For a visual overview of the area nowadays (with the 2001 NCDC land-cover database), see Figure 3.2.

Comprising more than seven million inhabitants according to the 2000 Census, the Chicago metropolitan area has greatly expanded since the 1960s (Table 3.1 and Figure 3.1), across several distinct basins and at the expense of the surrounding agricultural plains. Growth rates in 1960-2000 depict the speed of this suburban extension across several counties. They have to be compared to that of Cook County, where the city of Chicago itself lies. They suggest a progressive development farther and farther from downtown Chicago, and that the urbanization process in Cook County is virtually over by now, with a population stabilized between 5 and 5.5 million inhabitants. Table 3.1 also suggests a much denser urban habitat in and near the city of Chicago, as compared to a less dense suburban development in neighboring counties. The major hydrologic units that encompass the Greater Chicago are the DesPlaines River and the DuPage River. The former, closer to the lake, has undergone a longer urbanization than the latter (*Changnon and Demissie*, 1996). They both flow from North to South, and are where the aggregated effects of urban development can be best observed.

Rural basins serve as a means to compare the effects of a similar climate on watersheds that have drastically different land uses. South of Chicago, the agricultural Kankakee and Vermilion River basins were chosen. On the Western outskirts of Chicago is the Fox River that flows from Wisconsin, a predominantly rural basin but with an urbanizing midstream (*McConkey et al.*, 2004). For that reason two watersheds were also picked further West of Chicago. In total the outlets of 10 rural watersheds are studied, alongside with those of 26 urban ones, to reach a total of 36 watersheds.

3.2.2 Data

This study was conducted using daily streamflow and rainfall data, with records that are long enough to carry local meaningful analysis and pattern recognition. 40 years is considered to be a minimum for that purpose. Rainfall series are used to build a broad understanding of how climatic variability affects streamflow, thus helping us detect which change patterns in streamflow may have other causes than climate.

Daily streamflow data is the core of this analysis. It comprises all the 36 streamflow gauges with more than 40 years of records operated by USGS (United States Geological Survey) in the study area. They are presented in Table 3.2 and Figure 3.2, and classified by temporal span of the record, land use type and size of drainage area. The water year is chosen to be the calendar year because winter is the period with least incoming precipitation. This reduces the chances to change water year in the middle of a flow event. The first year reported in Table 3.2 is not the year when the gage started being operated, but it corresponds to the first complete year of operation. There is an exception for Weller Creek (gage 05530000): this gage displays a strong effluent discharge on the period 1951-1958 (as confirmed by *Meyer*, 2005). Consequently, 1959 was used as the first year of records at this particular location, not to have the streamflow decline caused by the end of the effluent discharge impeding any relevant trend analysis on posterior increases. For all gages, records are studied until December 31^{st} of 2008. One of the gages, though, has its records interrupted in 2008: DuPage River at Shorewood (05540500). Records from this gage are only considered until December 31^{st} of 2007.

One can note from Table 3.2 that the dataset is heavily biased towards urban gages. This is due to the absence of gauges for small rural watersheds (under $100km^2$). Such data could have provided a great means of comparison with the 16 small urban basins.

Climate data from NCDC (National Climate Data Center) is used to understand the impacts of rainfall and temperature (through evaporation) on streamflow. Divisional data over the whole northeastern Illinois is considered, as well as local rainfall data from rain gauges. Radar data is unavailable because it is too recent for our historical analysis. Finding rain gauges with long and relatively uninterrupted records also proved to be difficult. The selected rain gauges are summarized in Table 3.3 and represented in Figure 3.2. Only monthly data was used, to make inferences on rainfall amount trends. Their spatial repartition allows a better understanding of more local rainfall patterns, and eventually to try and compare rainfall and streamflow patterns. According to *Pryor et al.* (2009) the range of spatial decoherence of annual rainfall is of more than 200 km, which is much more than the resolution of our spatial gauge network (a few tens of kilometers). Besides rainfall data, monthly temperature data at the divisional level is also used to have an idea whether there is a relation between the changes in temperature and rainfall (*Lettenmaier et al.*, 1994; *Groisman et al.*, 2004).

3.2.3 Method for the analysis of streamflow

For each year, 38 indicators of streamflow are extracted from the daily records. The distribution of annual flow is studied through its deciles: the annual minimum (Q_0), the quantiles Q_{10}, Q_{20}, Q_{30}, Q_{40}, the median Q_{50}, and then again the quantiles Q_{60}, Q_{70}, Q_{80}, Q_{90}. The annual daily maximum, Q_{100} is also a variable of interest, as well as the 7-day minimum, the 7-day maximum and the mean annual flow. The mean as well as the quartiles (Q_0, Q_{25}, Q_{50}, Q_{75} and Q_{100}) of seasonal flow are extracted for each year, because seasonality of precipitation and evaporation is a major control over the water balance of a catchment (*Milly*, 1994b), while urbanization is known to affect evaporation (*Dow and DeWalle*, 2000). Fewer time-series are extracted from seasonal daily values than for annual ones because the sample is four times as small. The four seasons (Winter, Spring, Summer and Fall) are successively defined as 3-month periods during the calendar year.

For each of the series, the analysis is then conduced using the PCPS method described in section 2.4.3, and returns the following items:

- The sign of the MK statistics for the whole series. Do we have an increase or a decrease?
- The Pettitt year T, to give a first idea of the timing of change.
- For $X = 0.95 \times |S|$, $0.99 \times |S|$ and $|S|$, if $K < X$ we search for the period of change p_c introduced in equation (2.20). We also search for the statistical significance of $S(1, k_c)$ and $S(k_c + d_c + 1, n)$ for all levels of X.
- The significant autocorrelation coefficients ($\alpha = 0.1$) of the residuals after detrending, till lag 5.
- Most importantly, **whether change is statistically significant at the level** $\alpha = 0.05$.

- Whether a “strong trend” (as defined in section 2.4.4) exists.
- Whether this “strong trend” describes past change, depending on the correlation of the residuals of this trend (again, please refer to section 2.4.4).
- The last year y_e for which $S(y_e, n)$ has a sign that is not that of $S(1, n)$. It can be an indicator whether change is currently going on.

The same procedure is going to be applied to the rainfall time-series, to try and find inferences from precipitation to runoff. For that rainfall will be aggregated at the yearly and seasonal time scales.

The studied time-series are of different lengths. The possibility of making meaningful statistical analysis without the setting of a common period of record for all gauges like in many previous studies (e.g. *Lettenmaier et al.*, 1994; *Douglas et al.*, 2000; *Lins and Slack*, 1999, 2005) is an advantage of PCPS.

3.3 Change patterns for streamflow

This section presents the results of the analysis for streamflow. They are summarized in Table 3.4 and Figures 3.7, 3.8 and 3.9. All the significant shifts found over a whole time-series are flow increases, except for the low flows (until Q_{30}) of the Skokie River, a small urban catchment (stations 05535000 and 05535070). The section is organized so as to describe the difference between the patterns observed in urban and rural areas. Recall from the first part of this thesis (Chapter 2) that change patterns of very different durations of change observed in several gauges as the result of the application of PCPS are a sign of the presence of different underlying processes.

3.3.1 Rural areas: change during 1965-1972

Almost all the identified periods of change in Table 3.4 are between 1965 and 1972. This suggests that this region experienced a rather abrupt change, consistent with the date for a step change in streamflow in the eastern United States given by *McCabe and Wolock* (2002). Besides, most of the few strong trends detected in these basins display a change that is historical rather than ongoing. Change affects systematically the annual mean and low flows, but not always the 1 or 7-day maxima. This is consistent with nationwide studies (*Lins and Slack*, 1999, 2005). Fall is the most impacted season, even though spring and summer also experience some changes.

Spatio-temporal patterns in Figures 3.7 and 3.8 show that rural gauges with more than 50 years of records experience a step increase in low and mean flow in 1965-1972. The step increase in daily annual maximum is significant only in the southern half in the area (3.9). Yet, the 3 agricultural catchments with records starting after 1960 (less than 50 years) selected in this study exhibit no significant change (except for the winter low to median flows at Ferson Creek). This further suggests that change can be dated around 1965-1970, rather than after.

3.3.2 Urban areas: diverse patterns, long-term changes

The median of the years of change given for urban catchments in the Chicago metropolitan area also encompass the 1965-1972 period, but change is longer for most indicators. There is also a significant number of strong trends, especially for low and median flow series. This suggests that there may be change mechanisms specific to urbanizing areas, some of which might still be at work nowadays. At the aggregated spatial scale of larger catchments, summer and fall increases exist in all the chosen indicators, while only winter and spring peak flows show no significant change.

Figures 3.7, 3.8 and 3.9 show how temporal change patterns in streamflow in urban areas are more diverse than rural areas, but it also gives clues as to how they are spatially organized. Results for the annual daily maximum (Figure 3.9) are different from those for low and mean flow (3.8 and 3.7). But first, the significance of the 1965-1972 period is investigated.

A) Role of the 1965-1972 period

When applied to mean streamflow, the Pettit date (main year of change) lies in the period of 1965-1972 in 16 of the 26 urban gauges considered in this study. This suggests that the same factor that provoked streamflow increases in rural areas between 1965 and 1972 was a decisive factor for change to be statistically significant in many time-series from urban catchments (even if it was not the only factor). A close examination of the results shows that some urban catchments with records starting before 1960 have a period of change between 1965 and 1972, as can also be seen in . This common pattern over watersheds displaying very different evolutions of their land-use could be an incentive for searching a hydroclimatic cause to the step change. Temporal patterns in rainfall time-series will be examined in section 3.4.

B) Changes in peak flow

For maximum annual flow (Figure 3.9) only 15 of 36 (urban and rural) gauges show a statistically significant increase at $\alpha = 0.05$. Furthermore, the only 5 gauges that show possibly current change (strong trends) in that same indicator are situated on the edge of the suburbs, and concern only small watersheds. The spatial pattern within the urbanized area is heterogeneous: many urban gauges have increases temporally located outside of the 1965-1972 period, but others show no increase at all. In both cases that suggests other physical explanation than for rural areas. Besides, Table 3.4 reports increases in spring daily maximum in 5 rural watersheds but in none of the urban ones. These results seem counterintuitive because urbanization, understood as land-use change, is known to lead to a general enhancement of maximum annual flow. We will explain it in section 3.5.

C) Changes in low and mean flow

When it comes to low and mean flow (Figures 3.8 and 3.7 respectively), long-term change is observed in most Chicago suburbs. Strong trends can be detected more towards the center of the metropolitan area, in the older suburbs which have a longer history of urbanization. Superposing the two maps suggests that many of the strong trends in mean annual flow are in central watersheds where low flows increase are due to past change. If the mechanisms for change in both indicators are related, this may suggest that even though current change is a statistically valid model for these mean flow time-series, it might not be physically supported. More strong trends are observed for low flows than for mean flow. Thus, several currently urbanizing areas north, west and south of Chicago show increases in 7-day minimum that are possibly ongoing. Yet, for both low and mean flows, patterns are more diversified in these suburbs at the periphery of the city. There even are decreases in 7-day minimum in north of the city. Because of the outwards extension of the Chicago suburbs, peripheral suburbs may represent newer ones, as compared to the older ones lying closer to downtown Chicago. Section 3.6 will be about explaining this opposition between the homogeneity of increases in older suburbs and the heterogeneity of the results for newer ones.

When it comes to mean flow, another striking result has to do with the special role of the summer season. It is the one with the most gradual, long-lasting changes. Table 3.5 shows the periods of change, as well as the latest year k_e for which $S(k_e, n) \leq 0$. These are two possible indicators for the end of the change period. Results indicate that the increases in the mean flow for the fall season generally occurred in the past. Likewise, most increases in winter and spring seem to be over. This is not the case for summer increases: most of the changes in the mean annual flow that seem to have occurred recently are accompanied by similar changes in summer flow. In particular, this connection is systematic for the 7 gages that have a strong trend in mean annual flow, as shown in 3.10 and 3.11 for the 6 of them that have more than the 50 years of records.

3.4 The role of rainfall variability

Now that we described the main features of change in streamflow, we must understand which role climatic controls play in it. This section will first describe the systematic change patterns that exist in local and regional rainfall records, then understand which streamflow change patterns can be explained by them. It is also important to understand how seasonal and local rainfall variability may hinder the subsequent analysis of the effects of urbanization on streamflow. Finally, attention will be given to the frequency of streamflow peaks, in order to shed light on the pattern displayed in Figure 3.9.

3.4.1 A step change in annual rainfall

Application of the PCPS methodology to divisional rainfall shows a step change in 1965. Figure 3.3 shows this abrupt shift in the rainfall record: the difference between the mean annual rainfall in 1895-1964 and 1965-2008 is 104 millimeters of rainfall per year over the whole Northeastern Illinois climate division of NCDC. It uses cumulative deviations from the mean, which can be a means to investigate shifts in rainfall (e.g. *Buishand*, 1982). This shift corresponds to a 12% increase in annual rainfall over the whole area. Such a shift is confirmed by the observations by *Smith and Richman* (1993), who depict it as part of a larger shift in the whole state of Illinois. Examination of divisional data from two neighboring Midwestern States confirm that the climate was wetter in 1965-2008 as compared to 1895-2008. This is the case for example in Southeastern Wisconsin, at the Fox River headwaters, with an increase of 79 millimeters per year, or 10%. Similarly in Northwestern Indiana, where the upstream part of the Kankakee River basin is, precipitation was enhanced by 56 mm per year between the two periods.

But despite this regional pattern, local differences can be discovered. *Smith and Richman* (1993) show differences in the relative rainfall increase between the 1950-1967 and 1968-1985 periods, ranging from 6% to 15%. Likewise, while the analysis for the 8 gages pictured a wetter climate after 1965 than before, with increase of at least 3 inches (75 mm) per year, there are strong local discrepancies. Interestingly, only 4 of the 8 gages confirm that the rainfall increase is significant at the 95% confidence level. This may be due to the strong year-to-year variance of rainfall: the greater the variance, the more difficult for the MK test to detect trends (*Yue et al.*, 2002a). A sample of these are presented in Figure 3.4. Two rainfall gages, one South of Chicago (at Joliet) and one North of Chicago (at Waukegan), exhibit rainfall levels in the 1980s that are comparable to those of the period before 1965. Yet, all the other ones display a more uniform temporal shift.

Is this shift imputable to global anthropogenic climate change for sure? Probably not, for two main reasons. The first reason has to do with developments in hydrologic theory that enable stationary processes to fluctuate over long temporal scales. Indeed, it is acknowledged since the analysis carried out by *Hurst* (1951) on the historic records of the Nile flood levels (the Nilometer series), that natural hydroclimatic processes can fluctuate over the course of decades or centuries. These variations can be modeled using stationary stochastic processes with long-range memory, introduced in the field of hydrology by *Mandelbrot and Wallis* (1968). *Koutsoyiannis* (2006) argues that when we account for this long-term persistence in the data, the range of variations that can be accounted for as natural is in fact considerable. That is why in a hydroclimatic series, it is often difficult to separate what comes from long-term variability to what can be attributed to anthropogenically forced climate change (e.g. *Koutsoyiannis and Montanari*, 2007; *Milly et al.*, 2008). Likewise, it remains difficult to statistically prove the effects of a systematic climatic change on streamflow (*Kundzewicz et al.*, 2005; *Svensson et al.*, 2005; *Villarini et al.*, 2009a).

The second reason has to do with the used climatic data itself. Divisional rainfall and temperature data were aggregated at the decadal temporal scale, to better understand their long-term fluctuations. The results are displayed on Figures 3.5 for rainfall and 3.6 for temperature, with decades running on periods of the type 1849-1958, 1959-1968, etc... While *Groisman et al.* (2004) suggest a link between increases in both climatic indicators in the United States over the last 50 years, here they are largely uncorrelated. For example, while the temperature increases of the last thirty years could well be a local reflection of global warming, they are not accompanied by significant nor even meaningful changes in rainfall. Only studies at a much broader scale could relate these variations to global climate change. Further, it is difficult to make sense out of seasonal patterns, as they don't show any systematic pattern.

3.4.2 Regional response in streamflow

This step change in rainfall can be linked with the results described in section 3.3. Indeed, there we highlighted how most of the changes were recorded between 1965 and 1972, especially for rural areas. For the gages where 1965-1972 is the period of change, the coincidence of rainfall and streamflow increases suggests that hydroclimatic variability may be the major driver of streamflow change.

Streamflow gauges with more gradual changes are almost always situated in urban areas: this suggests urbanization produces effects which, in the case of Chicago, are more gradual than the rainfall fluctuations of the last 60 years. As expected, increases in mean annual runoff are observed because of this streamflow increase. Low flows are also affected, which is consistent with *Smith and Richman* (1993). Rural maximum annual flow is affected in large basins in the southern half of the study area (Figure 3.9). Consistently with our analysis of seasonal rainfall, increases in fall are quasi-systematic while they are much more scarce in winter and spring (Table 3.4). Yet, a closer examination of summer increases seems necessary, especially in urban areas (Table 3.5).

Thus, a comparison between rainfall and streamflow patterns suggests that climatic varibility triggered streamflow changes in rural areas and that comparatively longer change patterns in urban areas should be side-effects of land-use change. But to use rural basins as a valid means of comparison, we need to further understand how increases in rainfall and streamflow can be related. For example tile drainage is known to have played a role in enhancing runoff in the Midwest (*Changnon and Demissie*, 1996; *Kumar et al.*, 2009). A conceptual mass balance calculation can be performed to understand the relation between rainfall and runoff increases on the Fox River basin, a mainly agricultural watershed. During the 1915-1964 period, the annual flow at the Dayton gage at the outlet of the Fox River Basin is equivalent to a depth of water of 192mm over the entire watershed. During the 1965-2008 period, it became equivalent to a depth of water of 296mm. Meanwhile, assuming that the average rainfall North of the New Munster gage is the same as the NCDC divisional data for Southeastern Wisconsin, while the average rainfall on the rest of the watershed corresponds to the divisional data for Northeastern Illinois, averaged rainfall in the watershed was 824mm for 1915-1964 and 921mm for 1965-2008. As a result, we have an estimated 97mm of increase in rainfall and a measured 104mm increase in streamflow. Given the roughness of the assessment, the latter figures suggest that rainfall is the main explanatory factor for streamflow change.

3.4.3 Rainfall variability and seasonal and local streamflow patterns

For rainfall, we see how the shift in annual rainfall is only the aggregation of the recent fluctuations of seasonal rainfall, with peaks in 1969-1978 for spring rainfall, and in 1979-1988 for summer and fall rainfall. Only fall rainfall shows a real systematic increase, albeit much weaker than the yearly increase. This shift in fall rainfall locally comfirms the findings by *Small et al.* (2006) at the regional scale and by *Lettenmaier et al.* (1994) at the nationwide level. It is detected by the PCPS at the divisional level, and for the same gages for which annual rainfall change was statistically confirmed at the 5% level. No significant change is detected for the whole period for any of the other seasons.

Thus, seasonal variations complicate the analysis of the effects of urbanization on runoff. For example, fall rainfall peaked in the 1980s, and this corresponds to the dates of end of change for the fall mean of many urban basins as indicated in Table 3.5. Any changes due to urbanization after that period are offset by lower rainfall levels. In a similar way, the summers of 2007 and 2008 are two of the five wettest summers of record for Northeastern Illinois, so that their presence right at the end of the study period biases the ending dates in the data. For instance DuPage River at Shorewood (05540500) is a gage situated at the outlet of a large urbanizing watershed for which the records ends in 2007. Its change period runs from 1950 to 1996 only, which suggests that 2007 and 2008 summers really did bias the results presented in Table 3.5. Also Figure 3.5 shows a multidecadal increase in summer rainfall from 1939-1948 to 1979-1988, which contributes to long-term changes in summer flow.

Local rainfall variability may also hinder our understanding. For instance in Table 3.5, mean summer flow exhibit more recent changes for the southern gages than for the northern ones. This is explained by a decreasing summer rainfall in the northern half of the study area. Meanwhile, there is no clear mid- or long-term pattern in the southern half. Figure 3.12 shows the data for the four Northernmost rain gages. The linear trends plotted in red are from 1977 onwards, because 1977 is the year of (non-signficant) change for these four gages. One can see that summer rainfall follows a decreasing trend, North but also West of Chicago (where the Wheaton gage is). This explains the recent decreasing trends in 7-day minimum that were detected in Figure 3.8 for some of the northern suburbs. Yet, these decreases are not a rule, as one can see in this same figure for the Chicago River gage at Northbrook, at the outlet a small watershed parallel to Skokie River. They disappear at the water flow South in the DesPlaines River, from Gurnee to Des Plaines and then Riverside. Effluent discharge can easily offset the adverse consequences of decreased flow, as is the case for the Des Plaines River from 1978 onwards.

3.4.4 Rainfall and streamflow peaks

Trends in rainfall in the United States were extensively studied by *Karl and Knight* (1998), who primarily attribute upwards trends to an increase in frequency and magnitude of extreme events. These findings suggest that rainfall may enhance direct runoff more than infiltration or evaporation. They were confirmed by posterior studies like the one by *Pryor et al.* (2009) who pinpoint that the Upper Midwest and the Great Lakes region are the areas were this trend in extreme rainfall events is most observed. Finally, *Groisman et al.* (2001) relate the increase in heavy precipitations in the eastern United States with that of peak streamflow. This is precisely the effect that we want to investigate locally. For that we need more information than simply the magnitude contained in the annual maxima (AM) series. We need also some information on frequency.

An alternative sampling method, called peak-over-threshold (POT), can be used to generate a time-series to investigate streamflow change for that purpose (*Svensson et al.*, 2005; *Renard et al.*, 2006). It selects events which peak over the chosen threshold. Analysis can be carried out on the evolution of the frequency of the peaks as well as on that of their magnitude. The latter is derived from the partial duration series that documents the temporal lag between two peaks. *Svensson et al.* (2005) found that AM were more effective at finding trends than POT magnitude series. However, POT partial duration series provide additional information that AM series, by nature, cannot give. That is the reason why it is found to be more effective at characterizing flood seasons in the Yangtze River basin by *Liu et al.* (2010). Partial duration will be used here to investigate further the coincidence between the step changes in annual rainfall and mean flow on one hand, and annual maxima on the other hand.

Peaks are defined as maxima in a centered window of 15 days. In other words, the value on a day of peak must be greater than all the daily values within a week before or after. If two days within 7 days of each other have the same streamflow value and can both be selected as peaks, only the first event is selected. The aim of doing so is to have peaks corresponding to distinct rainfall events. The minimum spacing of 8 days between two peaks has been determined through looking closely at the time-series of the Kankakee River basin, which is the largest basin studied here. It is a conservative value as compared to what can be found in the literature for example for POT data, *Svensson et al.* (2005) use 5 days for basins of less than $45,000\,km^2$. And it is known that POT only concerns a small sub-sample of all the possible peaks.

For all rural basins we have a pattern similar to that of Figure 3.14 for thresholds for which we have roughly as many peaks as the number of years of records. When the count of peaks is linear, the probability of occurrence of the peaks can be considered stationary. A change in frequency around 1972 can then be related to the step change observed in other times-series at the same site. This indicates that not only the magnitude, but also the frequency of high flow events increased, further justifying our earlier findings. It also potentially sheds light on the trends in extreme events discovered in the literature, suggesting that they may in fact be the result of a shift, at least locally. Finally, it poses the question of the absence of increase in maximum annual flow for urban times-series (Figure 3.9), under the conjugated effects of land-use change and enhanced storm frequency.

3.5 Aggregated impacts on high flows

Now that we described the local effects of rainfall variability on different indicators of streamflow, we can try and determine the major impacts of urbanization at different scales. We assume that the systematic shifts in streamflow that cannot be linked to a climatic explanation may be attributed to urbanization. In this section we investigate the impacts on maximum annual flow, then in section 3.6 we will deal with those on low to medium flows. This distinction is due to the differences between the former and the two latter types of indicators, as highlighted in Figure 3.7 to 3.9 and in Table 3.4.

We saw in section 3.4.4 that the regional rainfall shift was accompanied by an increase in maximum annual flow. But most urban gages show an increase posterior to 1972 without a strong trend associated with it, or no increase at all (Figure 3.9). Besides, changes in spring daily maximum flow were more widespread in rural areas, which is contrary to the expected effects of land-use change (*Lazaro*, 1976). Further, most of the gages that display an increase in maximum annual flow over the period of record also have y_e set in the 1980s at the latest. All this suggests that there is a mechanism countering the impact of urbanization on maximum annual flow. This section relates the observed patterns to stormwater management facilities, in order to understand their city-scale impact.

The previous assessment of maximum annual flow tackled annual maxima alone. We are going to investigate trends in frequency to refine our analysis. This will highlight the case of Addison Creek, which will in turn shed light on the results displayed in Figure 3.9 across different spatial urban scales.

3.5.1 Peak frequency analysis

Like in section 3.4.4, this investigation on urban peak behavior is carried out using the POT method. For each watershed, the largest 10% of all peaks are selected. The study period starts on 1970 to avoid dealing with the effects of the rainfall shift on event frequency. This roughly corresponds to a little more than forty peaks for each gauge, depending on the watershed. In other words, this amounts to a little more than one peak per year. Due to the high number of identical peak values in the time-series, we use the resampling method described in section 2.2.4; it corresponds to the MK trend test. It is applied to all the gages considered in this study to understand the differences between urban and agricultural watersheds, as well as across distinct scales. Given that the period of record is short, the number of trends detected may be weak. However, the motivation here is not to come up with a rigorous assessment of which gages follow which pattern, but to have an idea of the kind of shift in frequency that might take place in the area independently of any significant shift in extreme rainfall event frequency. Also, the results are only presented per group of watersheds: agricultural (Figure 3.15), large urban (Figure 3.16) and small urban (Figure 3.17).

Trend analysis for rural watersheds confirms that no significant shift in frequency occurred in the area after 1970. Only Kankakee River displays a slight increase in frequency; this may be due to land-use changes. Interestingly, for large urban watersheds we have rather homogeneous results for most gages, and they are comparable to what happens in rural areas. Only the Shorewood gage at the outlet of the DuPage River basin, an urbanizing area West of Chicago, documents a significant increase in frequency. Like for maximum annual flow magnitude, watersheds displaying strong trends are on the outskirts of the metropolitan area. This is also the case for small urban watersheds displaying rising trends in frequency, such as Flag Creek, Poplar Creek and West Branch of DuPage River at Warrenville, which are situated in urbanizing areas West and Southwest of Chicago. For small urban watersheds, the results also appear to be much more heterogeneous. They don't follow any systematic geographic pattern, which explains the homogeneity observed when flow is aggregated at larger scales. This also rules out any inference of rainfall patterns in the results, except for very unlikely micro-climatic conditions. But the most interesting feature of Figure 3.17 is that a couple watersheds display a significantly decreasing trend in peak frequency. Here this is the case for Weller Creek and Addison Creek; both watersheds did not display any change in maximum annual flow over their period of record. The case of Addison Creek is studied more in detail now.

3.5.2 Impacts of detention basins: case-study

In fact, the statistically significant decreasing trend observed for Addison Creek can only be a direct consequence of the building of stormwater management facilities. In this basin, urban areas went up from covering 33% in the late 1950s to 91% in the late 1990s (*Hejazi and Markus*, 2009). These facilities are detention basins, and they are local and don't entail water transfers across watersheds. However, in other watersheds there may be significant transfers through stormwater sewers. It was possible to get the data for the capacity and building dates of the detention basins (Erik L. Gil, PE, personal communication). The two biggest reservoirs in the Addison Creek watershed were built in 1976 and in 1987. Together, they were enough to stop the trend towards an increase in annual maximum (left on Figure 3.18), and also to induce a decreasing trend in the frequency of high flow events. In addition, the reservoir built in 1976 coincides with a step increase in the magnitude and variance 7-day maximum (right on Figure 3.18), as it progressively releases the water that came in with the storm, with a speed depending on factors such as river level, weather forecast, or downstream flow. This adds to the increases in 7-day maximum due to land-use change (*Ferguson and Suckling*, 1990), even though these are detected in only very few watersheds (Table 3.4).

Planning and management of stormwater detention reservoirs can be understood as a constrained multiobjective optimization problem (*Yeh and Labadie*, 1997), so that all floods except the most extremes are kept below a reasonable threshold set by flood planning engineers. That explains why there is a significant decreasing trend in the frequency of the higher peaks in Addison Creek, but nothing detectable for the amplitude of the trends. Figure 3.18 suggests that the target is to keep the water level below a certain level except for some events for which the capacity threshold is exceeded. Thus the stormwater management facilities may have prevented flood events from becoming more frequent.

In Figure 3.19 the response to the September 2008 event is displayed. This event, which occurred between the 13^{th} and the 15^{th} of the month, corresponds to at least a 10-year storm for most parts of the Chicago metropolitan area (*Westcott*, 2009). Rainfall corresponds to the mean of the daily totals for the two gages of the ISWS Cook County precipitation gages network, gages 5 and 8. One can see how the recession process is interrupted by a phase where the detention basins are emptied, thus shifting upwards all the upper quantiles of the flow distribution. This is consistent with the description that *Solo-Gabriele and Perkins* (1997) gave of the recession process due to detention basins in another situation, in order to study their impact on sediment transport. Emptying the basin can take several days, possibly implying an enhanced evaporation, and/or leakage for a few days. In fact, while natural land can store some water till a certain threshold after which there is overland flow (*Chow et al.*, 1988), urban watersheds can store water in the detention basin. Thus flooding is a mere threshold process in both cases.

3.5.3 Impacts of detention basins across different scales

The case of Addison Creek seems to be an extreme one if we refer to the results of Figure 3.17. Nevertheless, almost all the basins don't necessarily show the decreases in flow return periods that one would expect from urbanization (e.g. *Hollis*, 1975; *Villarini et al.*, 2009b). This suggests that stormwater management facilities are present in most watersheds of the area, albeit with a locally heterogeneous efficiency. They are less present in some newly urbanizing watersheds, and these are the places where we see the expected increases in maximum annual flow magnitude and frequency.

The aggregated effects of these local ponds is seen in Figure 3.16. Large urban catchments have the same behavior as rural ones when it comes to the evolution of peak frequency. Once again, let us use the September 13-15 2008 event to understand the impact of aggregation at different scales. Figure 3.20 shows instantaneous data in September 2008 for three gages: Addison Creek (small watershed), which flows into Salt Creek just a few kilometers downstream of the Western Springs gage. Salt Creek then flows into Des Plaines River gauge, at the outlet of the biggest urban watershed under study. The red line is the average effluent discharge found in Salt Creek by *Wang* *and Cai* (2009). According to all the gages in northwestern Cook County as well as in DuPage County, there was no rainfall between September 15 and September 25, so that rainfall could not have disturbed the first 10 days of the recession process.

The operation of the reservoir in Addison Creek is obvious on this figure, but because of the routing properties of bigger catchments, it is not the same for the two bigger watersheds. Yet the recession curves do show abrupt changes in the slope of the recession curve. This is particularly obvious for Salt Creek: between September 20 and 25, the recession slows down at a level of flow (3mm/day) where recession was continuing for the two previous events. From equation (10) in the analysis by *Wang and Cai* (2009), we get closer to the level of effluent discharge when the recession curve significantly slows down. This suggests that at those dates, everything is as though effluent discharge were above 2mm/day instead of the 0.36mm/day that *Wang and Cai* found. It would be explained by detention basins being emptied. Besides, according to *Solo-Gabriele and Perkins* (1997), the slope of the recession curve is flattened just after the peak is over, because the basins start getting emptied. This can be observed in Figure 3.20 in the three September events in Salt Creek, and that the beginning of the recession for the Des Plaines River is well flattened. For Des Plaines River, however, further analyzing the beginning of the recession process is almost impossible without an intimate knowledge of the human-controlled stormwater management system. Indeed the processes are controlled by human planning, and detention basins are not the only possibility for stormwater management. For instance, a common stormwater sewers network, the TARP (Tunnel And Reservoir Plan), proved to be effective to reduce flooding in some parts of the greater Chicago area (*Butts and Shackleford*, 1992). In the end, Figure 3.20 shows how the flood peak is controlled at the level of bigger watersheds through control measures at the level of smaller catchments.

The construction of stormwater facilities in an urban context can be seen as an adaptive response (e.g. *Werner and McNamara*, 2007). It is motivated by the ever-increasing costs of flooding (e.g. *Katz et al.*, 2002). In the case of the Greater Chicago Area, this analysis suggests this adaptive response took place at large urban scales to counter the effects of land-use change. It shows that the final effect of urban development can be other than enhanced flooding, if socio-economic and institutional factors lead to the design of engineering responses.

3.6 Aggregated impacts on low-flows and mean flow

Due to the complexity of water fluxes in an urban context, it is not obvious what the causes of the change patterns observed in low and mean flow may be. For baseflow alone, there are many possible local human inferences, detailed in section 3.1. That is why a city-wide water balance is performed in this section, to better understand which of these fluxes may be important at the scale of the Chicago metropolitan area. Low flow patterns are discussed next, because low flow increases can obviously be a cause for mean flow increases.

3.6.1 A water balance

Our goal is to understand the contributions to streamflow at the aggregated level of the Chicago metropolitan area. For that we can extend the water balance by *Grimmond et al.* (1986) from a unit surface of horizontal soil to a large urbanized area. Here we can neglect direct water withdrawal from the streams connected to the city. In Chicago, surface water withdrawals come from Lake Michigan since the 1970s (e.g. *Wang and Cai*, 2009). We then take into account many distinct fluxes of water between the city, its rivers, and other bodies including its local shallow aquifers. The mass balance should include the soil layers in which significant water movements occur during the period of interest (*Grimmond et al.*, 1986). All the fluxes used in this water balance model are introduced in Figure 3.22 for any interval of time Δt. The observed data are streamflow data, which can be decomposed as:

$$Q = DR + B + Eff \tag{3.1}$$

where DR is direct runoff after a rainfall event, B is baseflow arising from interactions with shallow aquifers, and Eff is effluent discharge. The water balance for the city is written as:

$$W + R + P = Eff + DR + I + L + E \tag{3.2}$$

with the three sources of water being the rainfall R, local pumping P and water withdrawals impacting external water bodies, W. Here P only entails pumping from subsurface aquifers: pumping from deep bedrock formation amounts to a withdrawal from an external water body and is therefore included in W. The city loses water to the atmosphere via evaporation E, through a natural process and/or as a result of human activities (e.g industrial water use). Evaporation also includes interception. It also loses water to the local unconfined aquifer from soil infiltration I (due to rainfall or graden irrigation) or from leakage L from man-made water conveyors (combined sewers, etc.).

Now groundwater storage S is subject to the following balance equation:

$$\Delta S = (I + L) - (P + B) \tag{3.3}$$

Replacing equations (3.3) and (3.1) into (3.2) we can rewrite the formula from *Grimmond et al.* [1986], but this time at the scale of an entire city:

$$R + W = Q + E + \Delta S \tag{3.4}$$

The major difference with their formulation is that W is no more the piped-in water supply, which could come from interbasin transfers inside the city via pumping. It becomes the water that comes from external water bodies. A lumped water balance allows to suppress the local interbasin transfers, which can hinder the analysis at smaller spatial scales ((*Claessens et al.*, 2006).

We can decompose the three sources of water R, W and P into contributions to E, Q and S:

$$R = R_E + R_Q + R_S \tag{3.5}$$

$$W = W_E + W_Q + W_S \tag{3.6}$$

$$P = P_E + P_Q + P_S \tag{3.7}$$

and the contributions of the city to streamflow can be decomposed into the following terms:

$$DR = R_Q \tag{3.8}$$

$$Eff = W_Q + P_Q \tag{3.9}$$

but writing the baseflow term is complicated by the existence of pumping, and by the variations of soil storage in equation (3.4). Two assumptions are needed in order to eliminate them.

First let us restrain equation (3.4) to the case where the urbanization process is over. There is no more non-stationarity induced by land-use change. Let us then make the assumption commonly made when examining the water balance of natural catchments (e.g. *Gerrits et al.*, 2009), that changes in storage are negligible at the annual time scale. Then for Δt of over a year, we can assume that equation (3.4) can be written as:

$$Q = R + W - E \tag{3.10}$$

The second assumption is that $W >> P$: water is primarily withdrawn from outside of the city. This is the case in Chicago, especially in older suburbs enclosed in large urbanized areas. We can then write that in particular $W_E >> P_E$ so that evaporation can be understood as coming from two major sources:

$$E \approx W_E + R_E \tag{3.11}$$

and replacing equations (3.5), (3.6) and (3.11) into (3.10) leads to the master water balance equation for large urbanized areas where local shallow aquifer can only supply a negligible portion of the water supply:

$$Q = \underbrace{R_S + W_S}_{B} + \underbrace{W_Q}_{Eff} + \underbrace{R_Q}_{DR} \tag{3.12}$$

where the identification $Eff = W_Q$ is possible by putting $W_Q >> P_Q$ in equation (3.9). Baseflow comes from R_S and external water supply: W_S can come from leakage and/or garden irrigation in suburban areas (*Grimmond and Oke*, 1986). R_S is supposed to be reduced by urbanization due to the reduced amount of soil through which infiltration is possible.

We are now going to apply the water balance equation (3.12) to the Chicago area. We will combine it with the findings of the PCPS method to try and better understand the practical causes of change in low and mean flows.

3.6.2 On low flow increases

Figure 3.8 shows that the low-flow change pattern is consistent with the spatial outwards extension of the city. Gradual low flow increases have ceased in central suburbs, are ongoing farther out, while the periphery of the suburbs displays varied patterns of mixed increases and decreases. For instance in Figure 3.13, a drop in the 7-day minimum level is observed for Skokie River just before summer rainfall started decreasing in the area. In the outer suburbs, the assumption that withdrawals must come from outside does not hold, and the example of Skokie River suggests there may be interbasin transfers at the local level. This may also be due to pumping. The results reflect the facts that local studies in newly urbanizing watersheds may lead to very different results when it comes to low flow behavior.

We can retrieve the same kind of geographic change pattern by analyzing the 7-day minimum after 1970, as was performed for the frequency of flood peaks in section 3.5. The results of trend analysis are shown in Figure 3.21. They show no significant change in rural areas, and significant increases in many urban watersheds.

The homogeneous increases detected in older suburbs suggest the presence of a same set of phenomena at a larger scale. Besides, the outlets of the two largest urban basins (05532500 and 05540500) both show strong trends. They display past change at the outlet of DesPlaines River where urbanization is older (Figure 3.13.c), and possibly current change at that of DuPage River where it is newer. From equation (3.12), after 1970 low flow increases may be primarily due to W, through effluent discharge or baseflow. Effluent discharge through W_Q, for instance, is apparent in DesPlaines River from 1978 onwards (Figure 3.13.a to .c). Similarly, *Wang and Cai* examine the case of Salt Creek (Figure 2.14), located in the suburbs West of Chicago, where effluent discharge led to low flow increase as soon as water started to come from Lake Michigan instead of pumping from the local aquifer. That is, when large-scale urbanization led to the second assumption leading to equation (3.12).

In fact even taking into account different stages of urbanization, the impacts of flow regulation and effluent discharge alone are very widespread in the Chicago Area. In fact according to *Meyer* (2005), only 3 of the small urban watersheds investigated in this study (McDonald Creek, Weller Creek and Tinley Creek) are not affected by neither of these impacts. Yet the first two watersheds are in Northern suburbs where summer rainfall decreased in the last 30 years; thus, low flow increases cannot be explained by local nor regional rainfall patterns. Equation (3.12) shows that when annual rainfall is roughly constant, increases in baseflow itself could mainly come from external water inputs into the city, through leakage or garden irrigation. In fact *Meyer* still reports a baseflow increase during the low flow period from August to December for all three watersheds. Findings from PCPS seems to confirm this finding, as all 3 streams show strong trends for summer and fall low flows, as well as for annual Q_0, Q_10, Q_20 and 7-day minimum. For all of these strong trends, current change is a valid model. This suggests that in equation (3.12), water withdrawals from the city (W_S) taken alone can increase baseflow enough to offset any decrease in infiltration, at least during low flow months.

This suggests that external water withdrawals can lead to a rise of the groundwater table that increases low flows through baseflow even where there is no reported effluent discharge. They also increase the stage of a river through effluent discharge. This in turn may have an impact on the water table during low-flow months. It could imply a feedback of climate variability on effluent discharge, with water being transferred from the river to the aquifer during drier years. This might eventually be a reason why the daily minimum for Salt Creek or the DesPlaines River 7-day minimum exhibit a stronger year-to-year variability now that it is dominated by steady effluent discharge than when it was due to climate variability. Yet, this is not a conclusive evidence since the overall effects of external water withdrawals on the water table at large scale are beyond the scope of this paper.

The idea of a positive feedback of W_Q on the groundwater levels is supported by Figure 3.23. The Weller Creek gauge is in Des Plaines, where it flows into the DesPlaines River. A step increase in 7-day minimum is detected coincidentally with the one in DesPlaines River, also at a gauge in DesPlaines, (Figure 3.23). The step increase quadruples the low flows in the DesPlaines River, as could already be noticed for other gages in this River in Figure 3.13. Following the joint analysis of rainfall and streamflow, it is most likely due to a new and large source of effluent discharge, than to rainfall variability. The increase of low flow levels in Weller Creek could then be a consequence of the higher head in the DesPlaines River during the low flow season. Figure 3.13 is by no means a conclusive evidence, because we don't have the necessary information about near-surface permeability or about the groundwater levels in this exact area. Besides, the water transfer, which is less than $0.02m^3/s$, ends up flowing back in the DesPlaines River. But such transfers mean the end of no-flow days in Weller Creek for most years, which has beneficial impacts for ecology.

Finally, the water balance from section 3.6.1 sheds light on the increases documented in older suburbs by showing how demographic pressure leads to an adaptive response with water withdrawals, which offsets the effects of reduced infiltration. But it cannot explain the more diverse patterns found at the periphery of the urbanized area, because its assumptions (negligible withdrawals from shallow aquifers and local streams) don't apply anymore.

3.6.3 On mean flow increases

The several possible causes for the homogeneous increase in urban mean flow observed in Figure 3.7 can be discussed from equation (3.12). First, rainfall increase can impact both low and mean flows. This happened around 1970 in the area, but cannot explain the more gradual changes observed in urban areas in Table 3.4. Urbanization can have two types of effects. First through land-use change, runoff is enhanced by impervious surfaces, while evaporation and natural infiltration tend to be adversely affected. Second through an adaptive response to a rising water demand due to demographic pressure, total runoff is enhanced by low flow increases broadly due to water withdrawals from external water bodies.

Both effects are most felt during the summer season as 1) summer is the season with most evapotranspiration, so that suppressing vegetation could increase runoff, and 2) low flow increases affect mean discharge more during low flow months. For instance in McDonald Creek (Figure 3.24), where *Hejazi and Markus* (2009) report an increase of urbanized areas from 10% in 1954-1961 to 86% in 1996-1999, mean annual and summer flows show a similar behavior, and these gradual changes may reflect the gradual impact of urban development on both land-use and groundwater recharge via leakage and garden irrigation. Both can enter into the explanation for the long change period of urban summer discharge (Table 3.5). However the gradual increase in summer rainfall from the 1940s to the 1980s (Figure 3.5) and the wet summers in 2007-2008, dicussed in section 3.4.3, don't fully account for the differences between rural and urban areas. In fact, the effect of land-use change on evaporation could even offset streamflow depletion due to climatic factors such as increased temperature or decreased rainfall (*DeWalle et al.*, 2000). But city-specific or even local factors could come into play and counter this effect (*Dow and DeWalle*, 2000). For instance frequent summer rainfall in Chicago causes a stronger evaporation from built surfaces than the three other North-American cities considered in *Grimmond and Oke* (1995).

Unfortunately, it is difficult to separate the two possible human-induced causes of mean flow increase from the data. Indeed, flow regulation (*Meyer*, 2005) and detention basins (section 3.5) impede the use of usual baseflow separation techniques. Changes due to a reduction in evaporation only affect mean flow, and not the low flows, which might explain why the mean flow increase can be current in areas where low flows stopped increasing (Figure 3.8).

3.7 Summary and conclusions

Application of PCPS to hydroclimatic data in Northeastern Illinois demonstrated plausible connections between a shift in rainfall and a shift in streamflow. The results suggest that a step change taking place around 1965 is an important driver of streamflow increases out of Chicago. The observed climatic fluctuation cannot be related to climate change, but still seriously interferes with the analysis of urbanization-induced changes in streamflow. This is especially true for seasonal data, for which the rainfall series don't show any definite patterns, but only fluctuations that break the possibly gradual changes that result from urbanization. Local variability of rainfall is also a problem when it comes to understanding broader patterns. Yet the shift in annual rainfall allows to better separate and assess the effects of urbanization. Once the effects of climate variability are identified, it becomes possible to discuss the city-scale impacts of urbanization.

At the scale of a large city like Chicago, changes in low and mean flows are primarily driven by water transfers, while increases in maximum annual flow trigger a generalized engineering response to control them. Both mechanisms are likely to interact with the groundwater supply at a local scale, i.e. through leakage, thus resulting in a further enhancement of baseflow at large urban scales. They are also found to be relatively robust to low-intensity changes in rainfall such as those that happened during and after the 1970s, so that their signature in the data is not hindered by seasonal variability. Several human factors contribute to increases in mean flow, and they can be linked to land-use change and external water withdrawals.

These conclusions show that the basic effects of replacing natural land by impervious surfaces may not always be dominant. Stormwater facilities replace the soil in its storage role when high-intensity rainfall events occur, and that may also play a part in reducing runoff enhancement by urbanization. They store water, producing possibly significant leakage, as well as evaporation in the case of detention basins. Water transfers also play a part in increasing the available quantity of water that can infiltrate into the ground, whether through leakage, garden irrigation, or more importantly effluent discharge. These cumulated effects offset the consequences of a reduced natural infiltration. In fact, it seems that while the effects of recent urbanization in areas on the edge of the metropolitan area display locally varied and often complex patterns, the impacts of historical urbanization at the city scale are simpler and different. This eventually supports our initial motivation for getting a broader view of the impacts of urbanization at the city scale, instead of studying them at a local scale like in the literature.

3.8 Tables and figures

Counties	Population (×0.1M)			Population Increase (%)		
	1960	1980	2000	1960-1980	1980-2000	*1960-2000*
DuPage	3.1	6.6	9.0	110.2	37.2	*188.4*
Kane	2.1	2.8	4.0	33.7	45.2	*94.1*
Lake	2.9	4.4	6.4	50.0	46.3	*119.4*
McHenry	0.8	1.5	2.6	75.6	75.9	*208.8*
Will	1.9	3.2	5.0	69.3	54.8	*162.1*
Total suburban	10.9	18.5	27.1	69.5	46.8	*148.8*
Cook	51.3	52.5	53.8	2.4	2.3	*4.8*
Total metropolitan area	62.2	71.0	80.9	14.2	13.9	*30.1*

Table 3.1: Population growth rates over 20 year periods for several counties over Northeastern Illinois. County-wide census data.

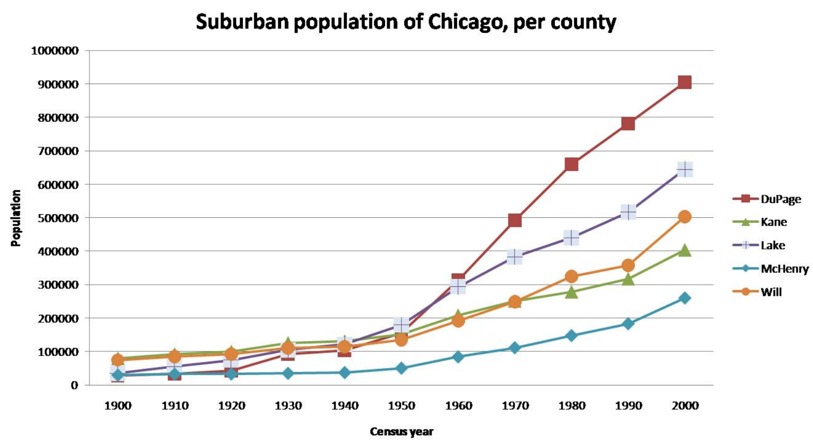

Figure 3.1: Census data for the 5 counties surrounding Cook county, where Chicago is located. Suburban development started in the post-war period to continue steadily until nowadays.

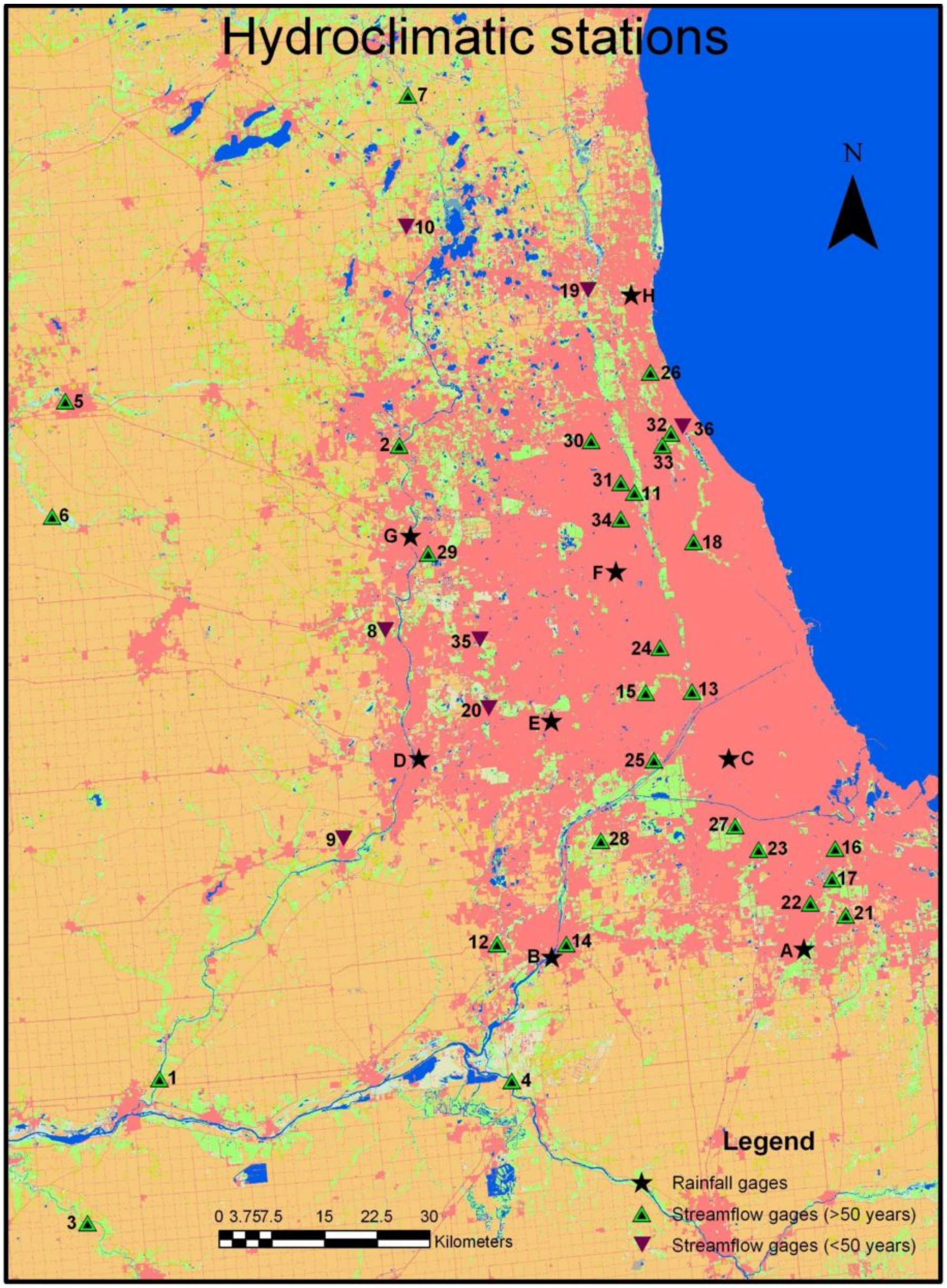

Figure 3.2: Map of rainfall and streamflow gages used. The background is land use: urban areas are shaded in light pink. Rainfall gages have the same letter as in Table 3.3, streamflow gages the same number as in Table 3.2.

		River	Site	Gage number	Drainage area (km^2)	Record starts in
Agricultural	1	Fox River	Dayton	05552500	6843	1915
	2	Fox River	Algonquin	05550000	3634	1916
	3	Vermilion River	Leonore	05555300	3240	1932
	4	Kankakee River	Wilmington	05527500	13338	1936
	5	Kishwaukee River	Belvidere	05438500	1393	1940
	6	SB Kishwaukee River	Fairdale	05439500	1002	1940
	7	Fox River	New Munster	05545750	2100	1940
	8	Ferson Creek	St Charles	05551200	134	1961
	9	Blackberry Creek	Yorkvile	05551700	182	1961
	10	Nippersink Creek	Spring Grove	05548280	497	1967
Large Urban	11	Des Plaines River	Des Plaines	05529000	932	1941
	12	Du Page River	Shorewood	05540500	839	1941
	13	Des Plaines River	Riverside	05532500	1631	1944
	14	Hickory Creek	Joliet	05539000	278	1945
	15	Salt Creek	Western Springs	05531500	298	1946
	16	Little Calumet River	South Holland	05536290	539	1948
	17	Thorn Creek	Thornton	05536275	269	1949
	18	NB Chicago River	Niles	05536000	259	1951
	19	Des Plaines River	Gurnee	05528000	601	1969
	20	WB Du Page River	Warrenville	05540095	234	1969
Small Urban	21	Deer Creek	Chicago Heights	05536235	60	1949
	22	Butterfield Creek	Flossmoor	05536255	61	1949
	23	Midlothian Creek	Oak Forest	05536340	33	1951
	24	Addison Creek	Bellwood	05532000	46	1952
	25	Flag Creek	Willow Springs	05533000	43	1952
	26	Skokie River	Lake Forest	05535000	34	1952
	27	Tinley Creek	Palos Park	05536500	29	1952
	28	Long Run	Lemont	05537500	54	1952
	29	Poplar Creek	Elgin	05550500	91	1952
	30	Buffalo Creek	Wheeling	05528500	51	1953
	31	McDonald Creek	Mount Prospect	05529500	21	1953
	32	NB Chicago River	Deerfield	05534500	51	1953
	33	WFNB Chicago River	Northbrook	05535500	30	1953
	34	Weller Creek	Des Plaines	05530000	34	1959
	35	WB Du Page River	West Chicago	05539900	74	1962
	36	Skokie River	Highland Park	05535070	57	1968

Table 3.2: List of the USGS streamflow gages with continuous records used for this study. They are classified by land-use type, first year of record used in this study, and gage number. They are numbered to facilitate their location in Figure 3.2.

NOAA ID		Station name	Latitude	Longitude	Period of record
116616	A	Park Forest	41°29'	-87°40'	1953-2008
114530	B	Joliet Brandom Rd Dam	41°30'	-88°06'	1941-2008
111577	C	Chicago Midway AP 3SW	41°44'	-87°46'	1942-2008
110338	D	Aurora	41°46'	-88°18'	1900-2008
119221	E	Wheaton 3 SE	41°48'	-88°04'	1937-2006
111549	F	Chicago O'Hare Intl. Ap.	41°59'	-87°56'	1959-2008
112736	G	Elgin	42°03'	-88°17'	1931-2008
119029	H	Waukegan	42°20'	-87°52'	1931-2001

Table 3.3: List of the NOAA rainfall gages used in this study. Letters are to locate them on the map 3.2.

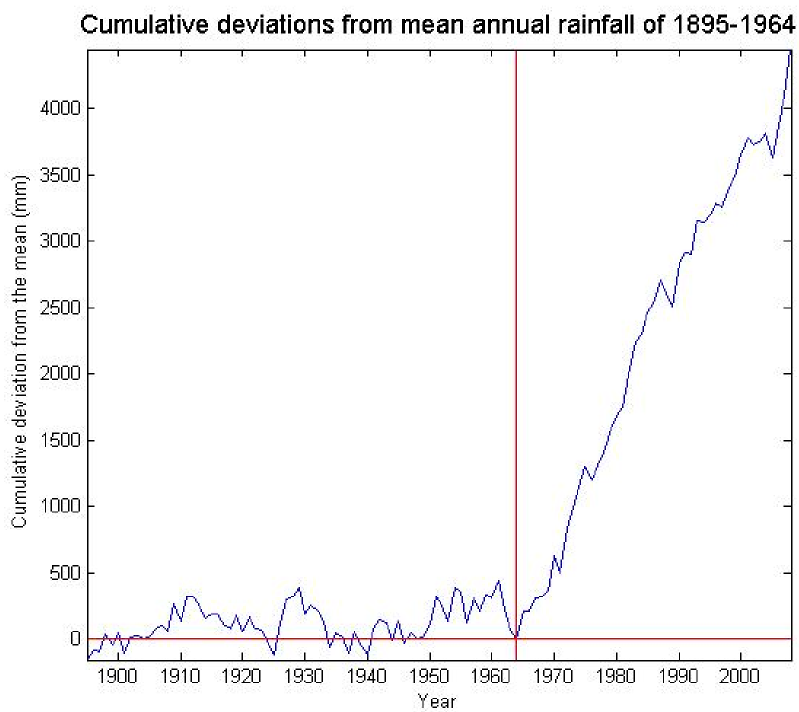

Figure 3.3: A shift in rainfall over Northeastern Illinois occurred after 1964. The mean annual rainfall is greater by 12.4% during $1965 - 2008$ as compared to $1895 - 1964$.

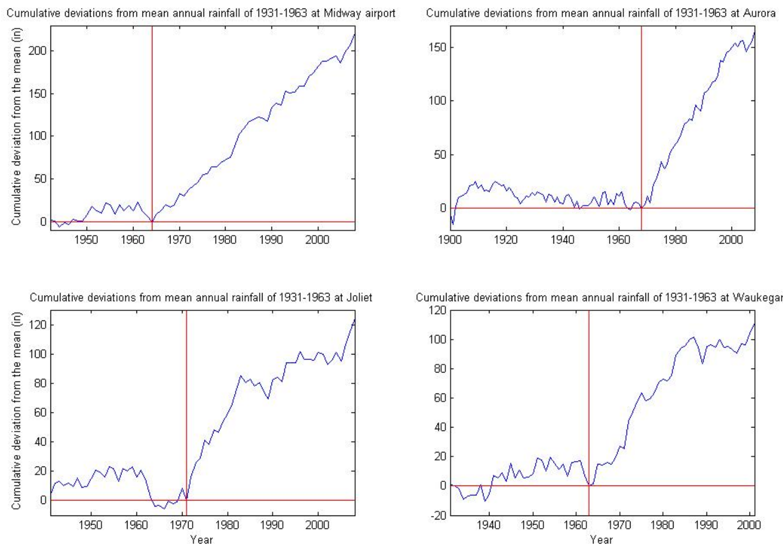

Figure 3.4: The two top figures show the same basic pattern as in Figure 3.3. The two bottom figures are the only 2 out of 8 gages that display a different behavior. They show that the regional shift doesn't prevent marked local pattern.

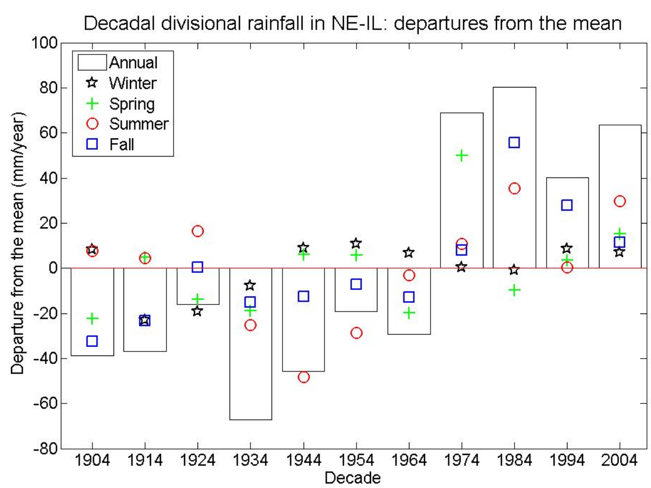

Figure 3.5: Deviations from the mean of 1899-2008 for decadal divisional rainfall quantities; the decades are 1899-1908, . . . till 1999-2008. The marked annual shift is a consequence of seasonal fluctuations that don't really describe marked shifts.

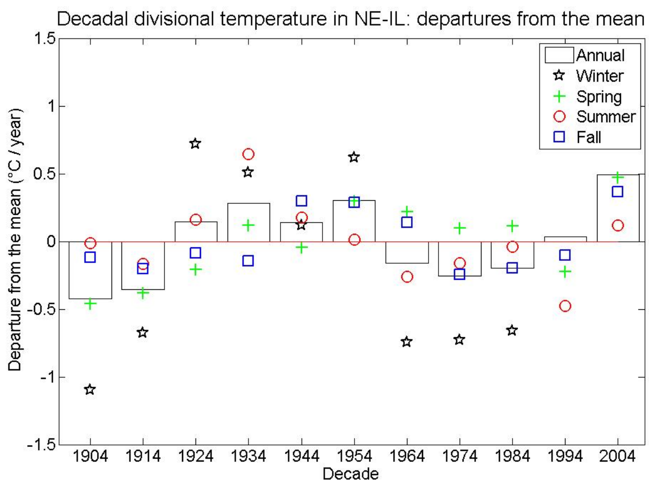

Figure 3.6: Deviations from the mean of 1899-2008 for decadal divisional temperatures; the decades are 1899-1908, . . . till 1999-2008. There is no marked pattern, but winter temperatures fluctuate importantly, with a difference of more than 3F (1.7°C) between 1969-1988 and 1989-2008.

	Indicator	Agricultural			Large urban			Small urban		
		# T	p_c	# st	# T	p_c	# st	# T	p_c	# st
Year-long	**Mean**	7	1969-1970	0	8	1965-1980	3	12	1965-1990	4
	Q_0	6	1967-1967	1(1)	10	1966-1982	7(2)	13	1969-1990	8(2)
	7-day min	7	1968-1968	0	9	1965-1981	7(3)	13	1968-1990	8(1)
	Q_{10}	8	1968-1972	0	10	1966-1984	7(2)	15	1967-1990	7(3)
	Q_{20}	7	1968-1968	0	10	1966-1985	5	15	1969-1990	8(1)
	Q_{30}	7	1968-1969	0	9	1965-1987	6	14	1967-1985	6(1)
	Q_{40}	7	1967-1967	0	9	1965-1988	4	13	1967-1989	5(1)
	Q_{50}	7	1967-1969	0	8	1965-1988	4	13	1970-1990	4(1)
	Q_{60}	6	1969-1969	1(1)	8	1965-1984	3	13	1969-1990	4
	Q_{70}	6	1969-1970	2(2)	8	1965-1980	3	12	1969-1989	4
	Q_{80}	6	1971-1971	0	7	1965-1978	3	12	1965-1972	3
	Q_{90}	6	1971-1971	0	7	1966-1972	2	13	1967-1973	4
	7-day max	3	1972-1972	0	2	1982-1982	0	6	1969-1982	1
	Q_{100}	4	1966-1968	0	6	1978-1982	1	11	1972-1982	4(1)
Winter	**Mean**	2	1969-1973	0	6	1973-1976	0	7	1973-1982	0
	Q_0	7	1969-1973	0	10	1966-1985	8(1)	12	1970-1987	4
	Q_{25}	7	1966-1973	0	9	1973-1986	5	12	1973-1987	4(1)
	Q_{50}	7	1973-1973	0	7	1973-1990	4(1)	10	1973-1989	4(2)
	Q_{75}	1	1973-1973	0	5	1971-1973	0	5	1965-1990	4
	Q_{100}	1	1973-1973	0	0	N/A	0	10	1974-1983	0
Spring	**Mean**	6	1965-1966	1	5	1965-1969	0	4	1969-1977	0
	Q_0	5	1966-1969	1(1)	6	1966-1979	3	9	1969-1982	1(1)
	Q_{25}	6	1969-1969	0	6	1966-1969	1	3	1965-1995	1
	Q_{50}	5	1966-1969	1	5	1969-1970	0	3	1966-1989	1
	Q_{75}	5	1965-1969	0	4	1965-1975	0	4	1969-1992	0
	Q_{100}	5	1965-1965	0	0	N/A	0	0	N/A	0
Summer	**Mean**	6	1968-1968	0	8	1954-1994	5	15	1967-2000	8
	Q_0	6	1967-1967	0	9	1968-1978	6(2)	13	1969-1990	8(1)
	Q_{25}	6	1968-1968	0	9	1964-1978	5	15	1967-1986	8(1)
	Q_{50}	6	1968-1968	0	8	1957-1978	6	13	1968-1990	7(1)
	Q_{75}	6	1968-1968	0	8	1957-1985	4	15	1963-1993	7
	Q_{100}	4	1958-1978	1	9	1957-2003	6	14	1968-1993	5(1)
Fall	**Mean**	7	1965-1965	0	8	1969-1981	1	13	1972-1982	1
	Q_0	6	1969-1970	0	9	1965-1985	6(1)	14	1968-1979	5
	Q_{25}	7	1969-1969	0	8	1966-1982	2	13	1967-1984	3
	Q_{50}	7	1967-1967	0	8	1967-1980	2	14	1971-1982	2
	Q_{75}	7	1967-1967	0	8	1970-1981	1	14	1972-1982	2
	Q_{100}	6	1965-1965	0	8	1966-1979	0	14	1970-1982	0
Sample size			**10**			**10**			**16**	

Table 3.4: Summary of the results found when applying the CDP method to the 38 time-series extracted from all 36 streamflow gauges. #T is the number of changes ($\alpha = 0.05$). #st is how many of them strong trend, and the parenthesis say how many of these display past change. Here p_c is the median period of change.

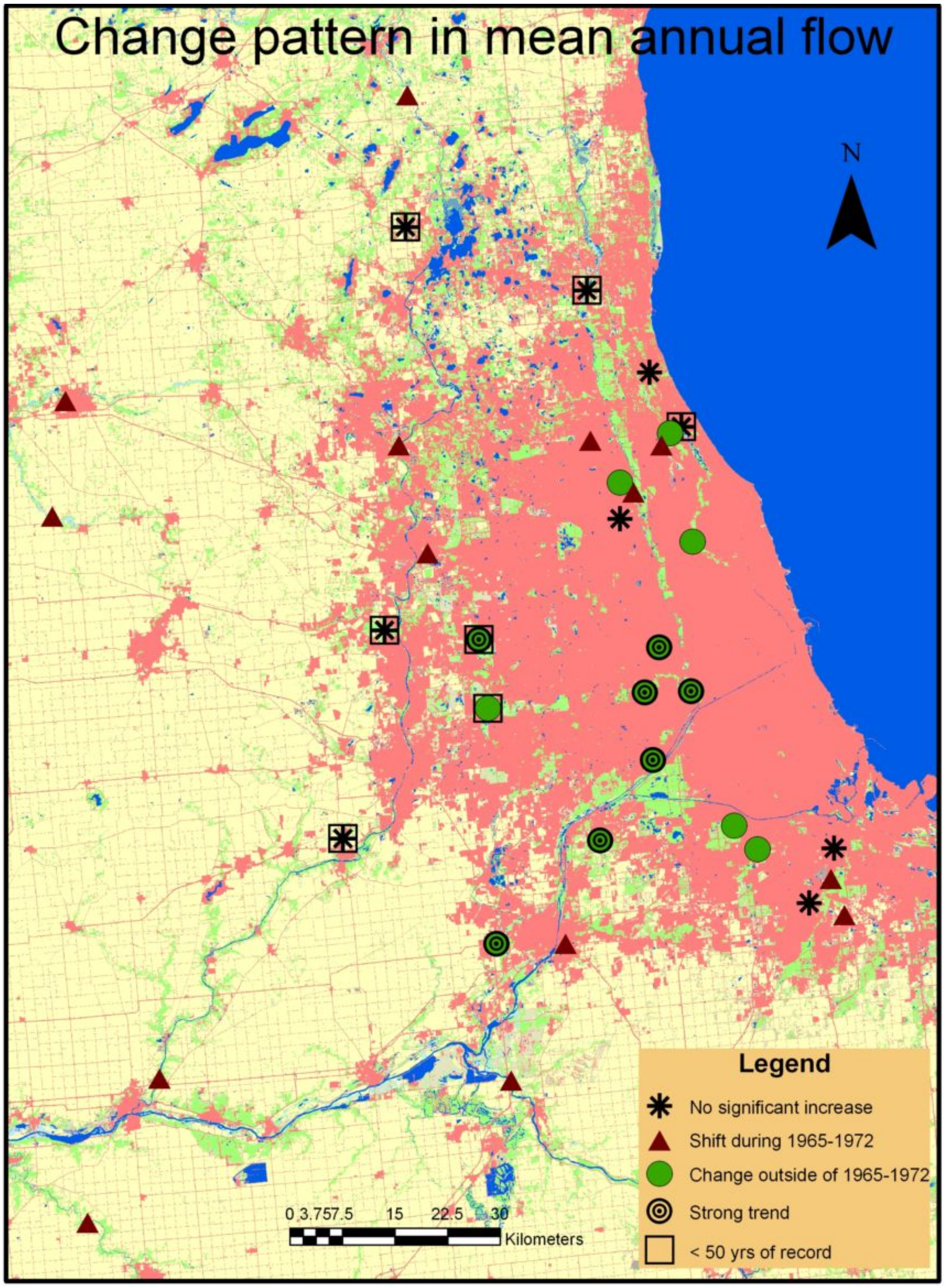

Figure 3.7: The shift in rainfall appears to be the main factor leading to streamflow increase in many cases. But distinct patterns emerge in some urban areas.

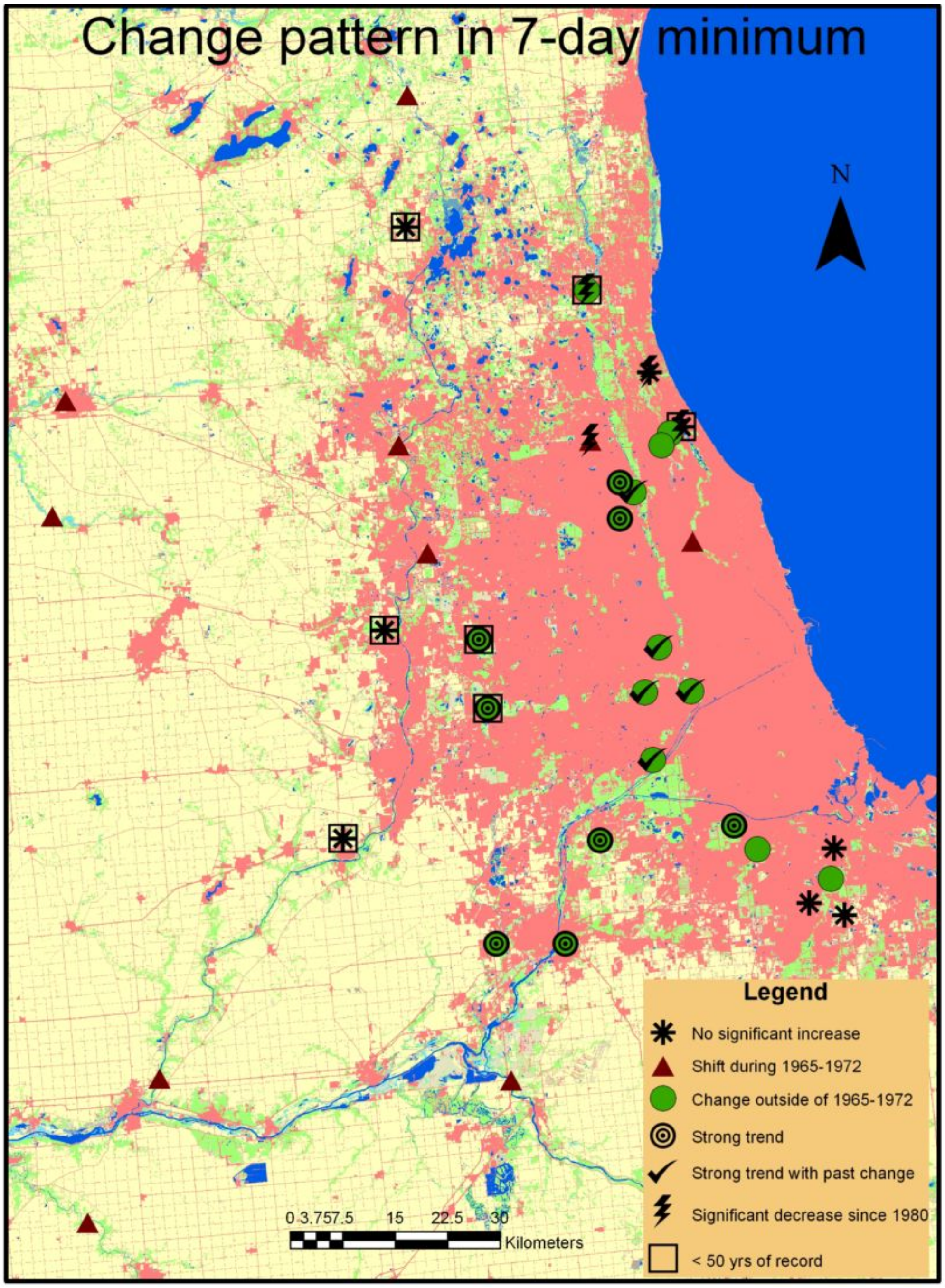

Figure 3.8: The general increase of low flows is a step change in rural areas, while urban areas exhibit a varied pattern with some long-term increases, but also some decreasing trends in Lake county.

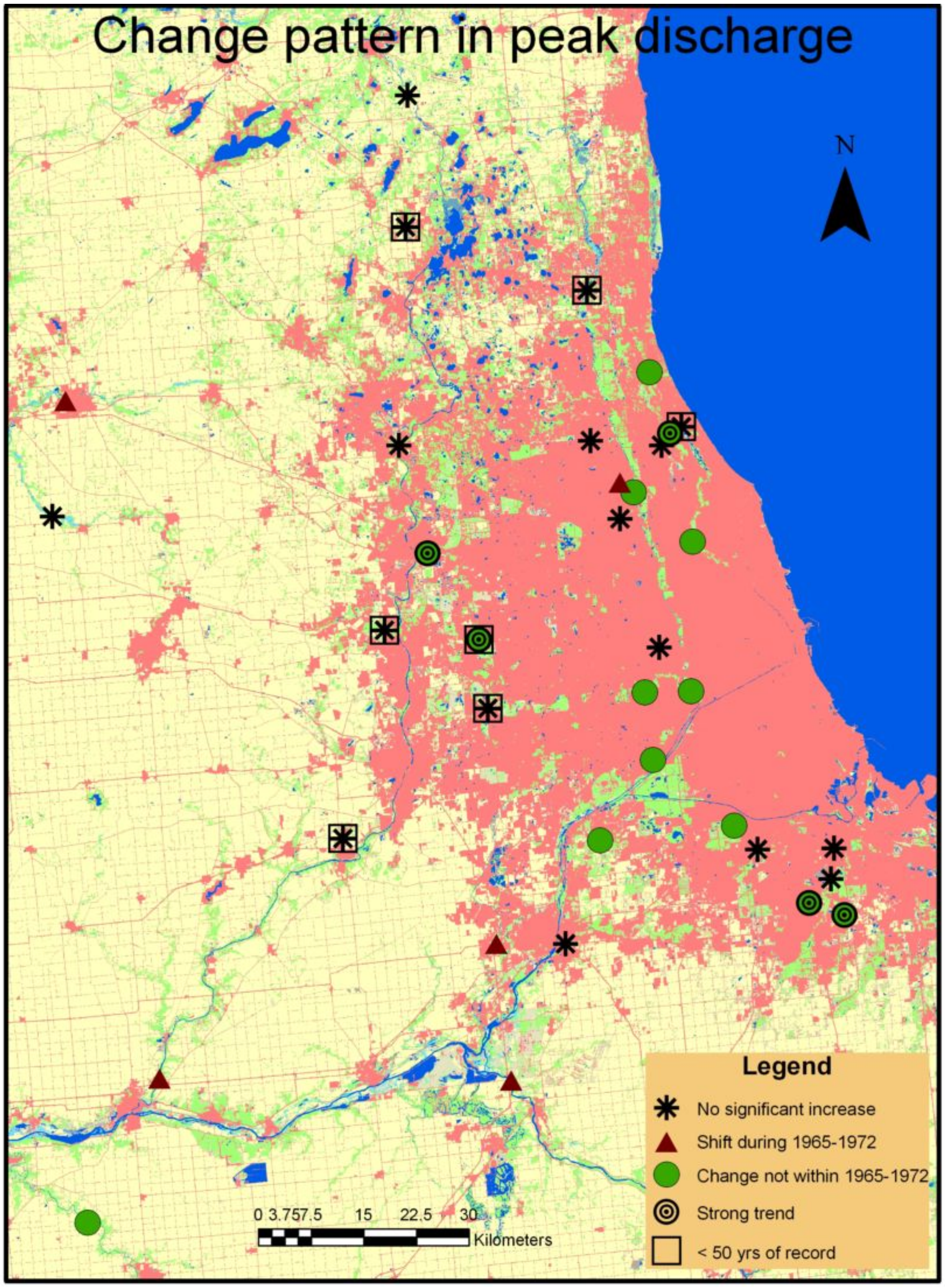

Figure 3.9: Peak discharge increase is frequent, but seems to belong to the past in most urban areas.

	Gage #	Annual mean		Winter		Spring		Summer		Fall	
		p_c	y_e	p_c	y_e	p_c	y_e	p_c	y_e	p_c	y_e
Agricultural	05552500	1969-1969	1965	1973-1973	1969	1965-1966	1965	1951-1965	1967	1965-1965	1969
	05550000	1969-1969	1971	–	–	1969-1969	1996	1968-1968	1967	1969-1969	1968
	05555300	1970-1970	1965	–	–	–	–	–	–	1965-1965	1965
	05527500	1968-1972	1971	1965-1973	1982	1966-1966	1964	1968-1986	1988	1965-1969	1977
	05438500	1972-1972	1969	–	–	1965-1965	1969	1968-1968	1967	1965-1965	1965
	05439500	1969-1969	1968	–	–	1965-1967	1966	1968-1968	1965	1965-1965	1964
	05545750	1972-1972	1969	–	–	1950-1972	1972	1968-1968	1967	1970-1970	1968
	05551200	–	–	–	–	–	–	–	–	–	–
	05551700	–	–	–	–	–	–	–	–	–	–
	05548280	–	–	–	–	–	–	–	–	–	–
Large Urban	05529000	1970-1972	1972	1973-1973	1993	1965-1969	1990	1964-1978	1977	1972-1978	1976
	05540500	1965-1989	1993	1973-1973	1990	1965-1969	1990	1950-1996	1996	1970-1982	1980
	05532500	1966-1982	1982	1973-1979	1997	1965-1969	1990	1951-1978	1978	1969-1980	1977
	05539000	1972-1972	2008	–	–	–	–	1976-1992	2008	1982-1982	1979
	05531500	1965-1990	1998	1973-1979	1990	1965-1969	1989	1947-2008	2008	1966-1982	1980
	05536290	–	–	–	–	–	–	1949-2008	2008	1965-1965	1964
	05536275	1965-1970	2008	–	–	–	–	1957-2006	2008	1972-1982	1979
	05536000	1965-1979	2008	1971-1973	1993	1965-1965	1990	1969-1977	1977	1968-1970	1970
	05528000	–	–	–	–	–	–	–	–	–	–
	05540095	1990-1990	1993	1982- 1993	1993	–	–	–	–	–	–
Small Urban	05536235	1966-1972	1990	–	–	–	–	1957-2005	2007	1965-1965	1965
	05536255	–	–	–	–	–	–	1968-2006	2008	1965-1965	1965
	05536340	1973-1990	2008	–	–	–	–	1968-1996	2008	1977-1982	1979
	05532000	1965-1997	2008	1965-1982	1990	1966-1966	1966	1956-2008	2008	1971-1982	1977
	05533000	1972-1990	1989	1973-1982	1982	1973-1989	1988	1957-2006	2008	1968-1982	1980
	05535000	–	–	–	–	–	–	1959-2000	2000	–	–
	05536500	1970-1982	1987	1973-1973	1973	–	–	1961-1996	1996	1979-1982	1979
	05537500	1966-1990	2008	–	–	–	–	1969-1996	2007	1982-1982	1977
	05550500	1965-1972	1990	1973-1973	1973	–	–	1969-1978	1977	1972-1980	1977
	05528500	1965-1972	1969	–	–	–	–	1972-1978	1977	1972-1982	1977
	05529500	1965-2007	2008	–	–	–	–	1957-2008	2008	1971-1982	1979
	05534500	1966-1990	1990	1965-1992	1990	–	–	1958-2000	2000	1972-1980	1977
	05535500	1965-1972	1972	1958-1997	2007	1965-1965	1964	1969-1978	1977	1964-1968	1968
	05530000	–	–	–	–	–	–	1968-2000	2000	–	–
	05539900	1965-2007	2008	1973-1990	1990	1990-1990	1990	1967-2007	2008	1977-1982	1978
	05535070	–	–	–	–	–	–	–	–	–	–

Table 3.5: While annual increases in the mean can mainly be related to summer and fall increases, changes in fall increases seem to be over while indications abund that summer increases may still be going on in urban areas.

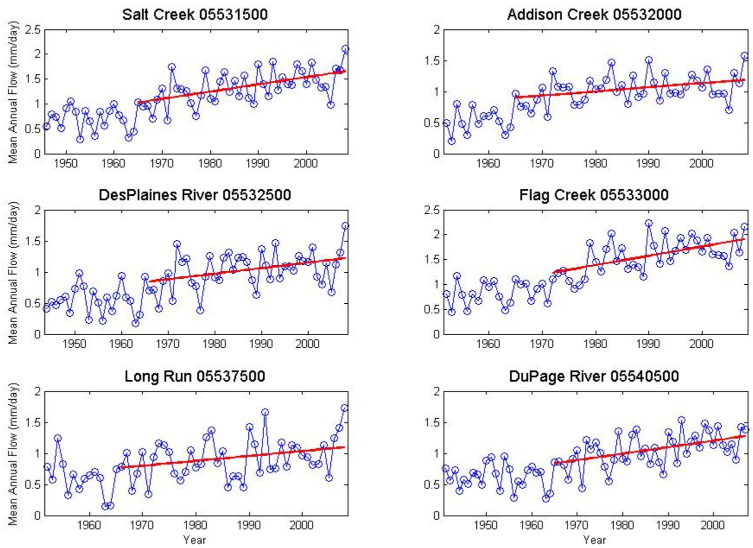

Figure 3.10: Annual mean flow at the gages for which a strong trend exists. Sometimes the trend is unconvincing, like for b) Addison Creek, perhaps because due to its closeness to Chicago, its urbanization is now a historical fact.

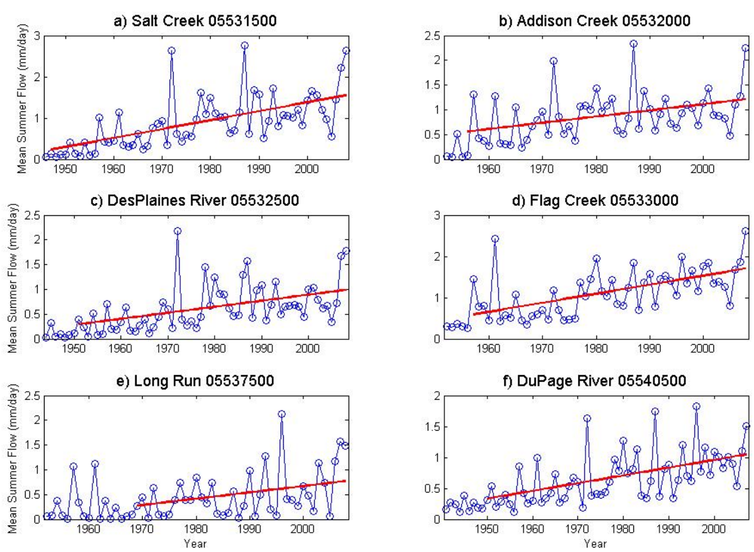

Figure 3.11: Mean summer flow at the gages for which a strong trend exists.

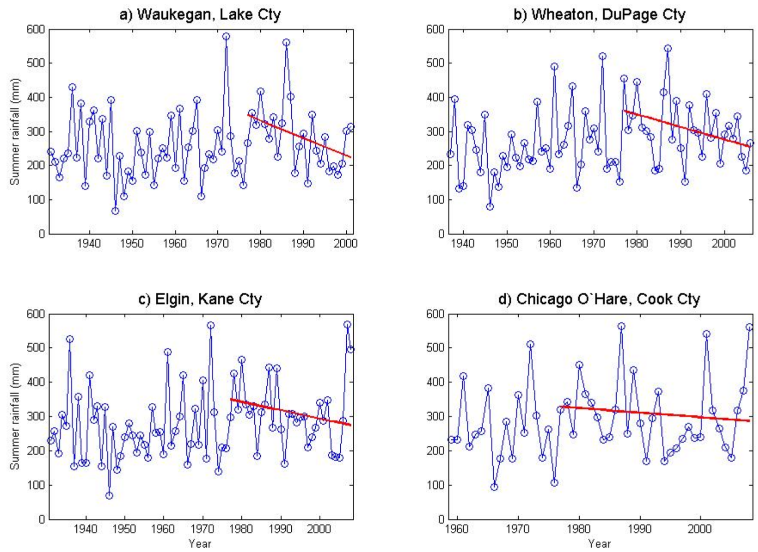

Figure 3.12: One can see how rainfall amount decreases since the wet 1977 summer, except at the O'Hare gage.

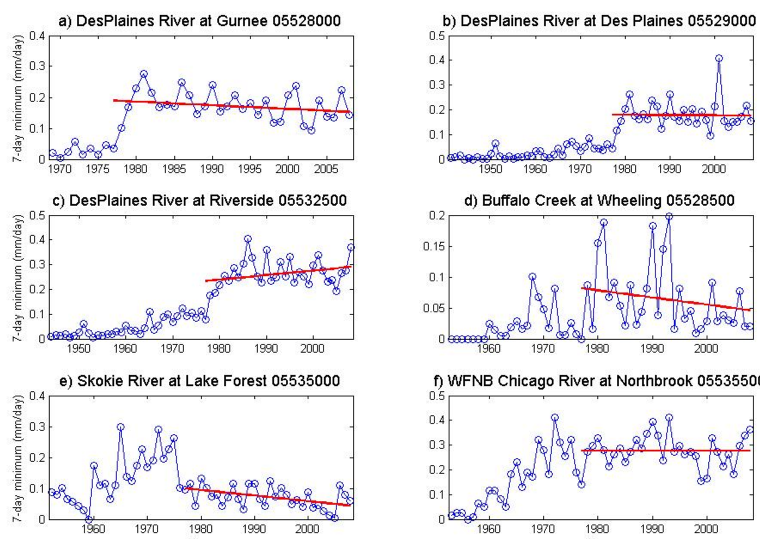

Figure 3.13: Some gages in the area show decreasing trends in the 7-day low flows consecutive to decreasing summer rainfall. But it is not a general pattern.

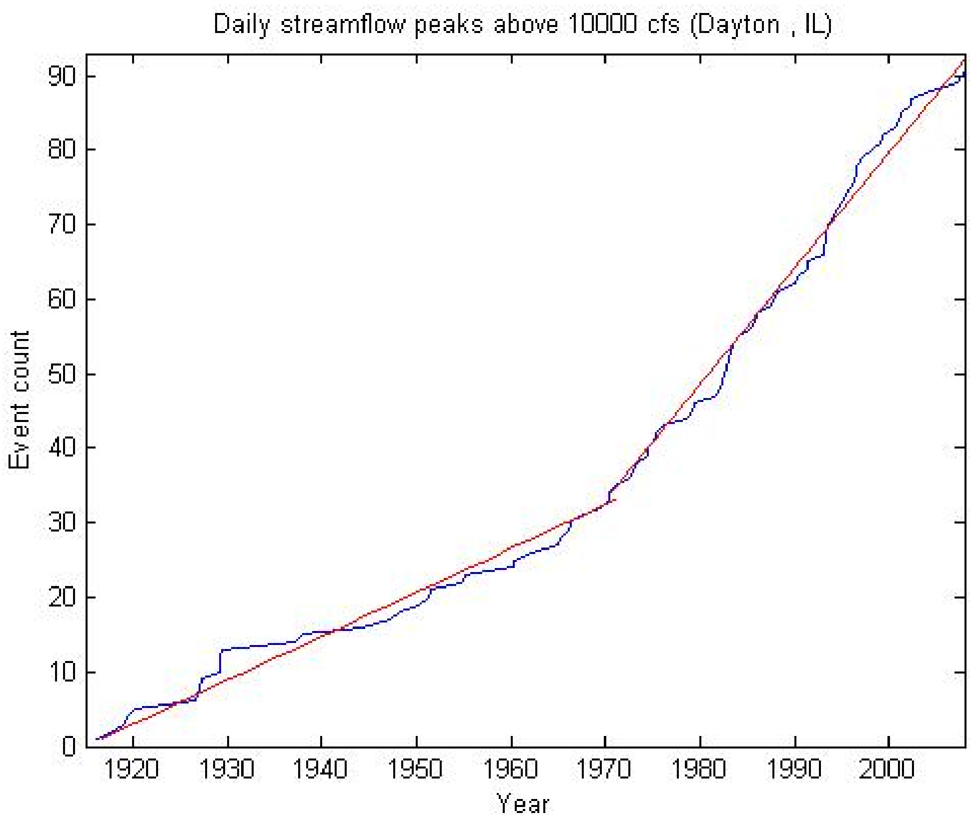

Figure 3.14: Peaks over 10,000 cfs are cumulatively counted over time at the outlet of the Fox River watershed. A constant slope can be interpreted as hydrologic stationarity. A break can be observed around 1972.

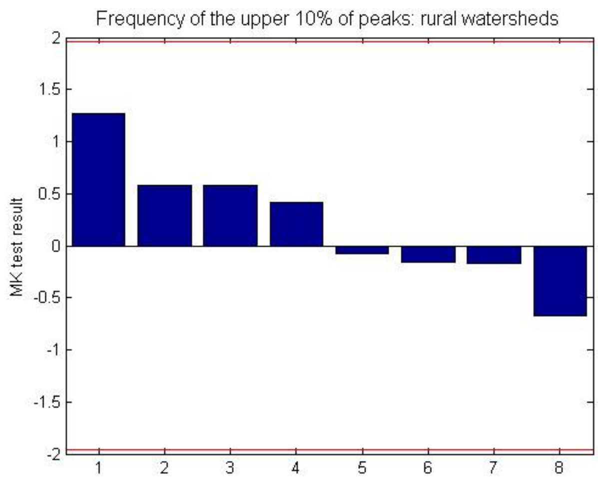

Figure 3.15: Trend test results for the evolution of the frequency of the 10% higher peaks for agricultural watersheds. The red lines mark the 5% significance level.

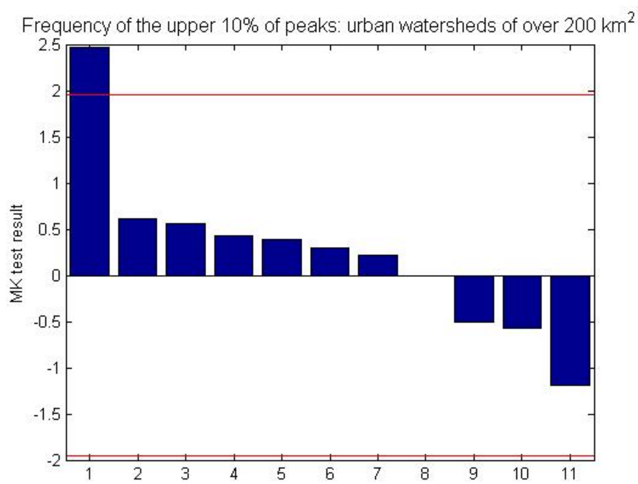

Figure 3.16: Same as above, but for large urban watersheds. Only the DuPage river basin, mainly covering newly urbanized areas, shows a significant trend.

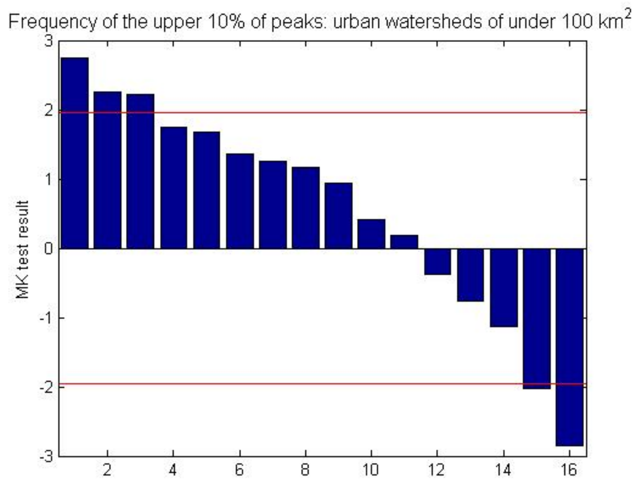

Figure 3.17: Same as above, but for small urban watersheds. Watersheds that show an increase are further away from Chicago that those that display a significant decrease.

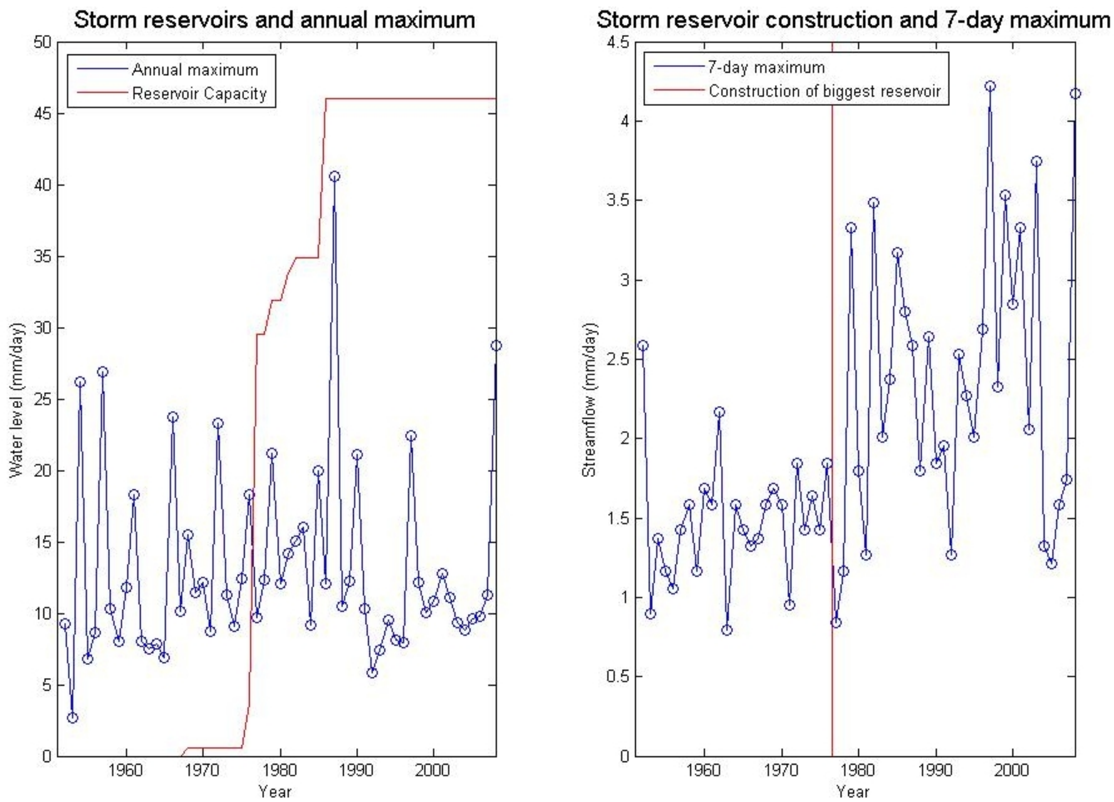

Figure 3.18: Left: evolution of annual maximum and reservoir capacity. Capacity is measured in height of water over the whole watershed. Right: impact of the biggest reservoir on 7-day maximum. Reservoir data is courtesy of Mr. Erik L. Gil, PE.

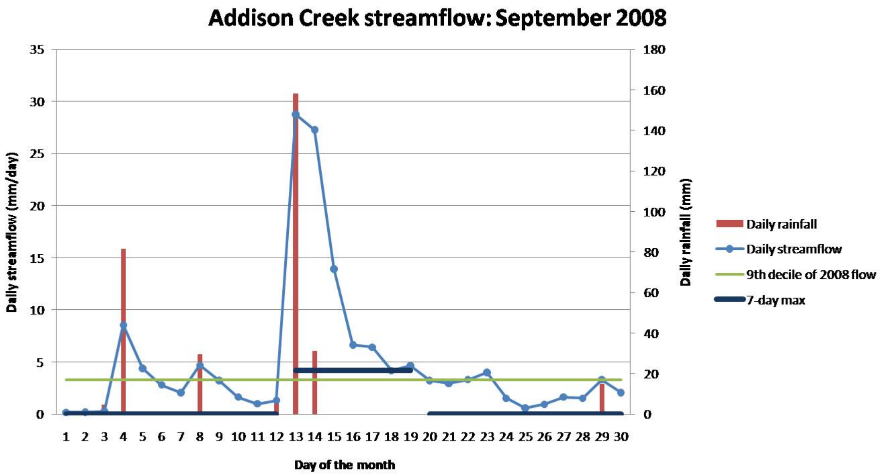

Figure 3.19: Response of Addison Creek to the September 2008 record rainfall. The impact of the offstream reservoirs on the recession process is clearly visible.

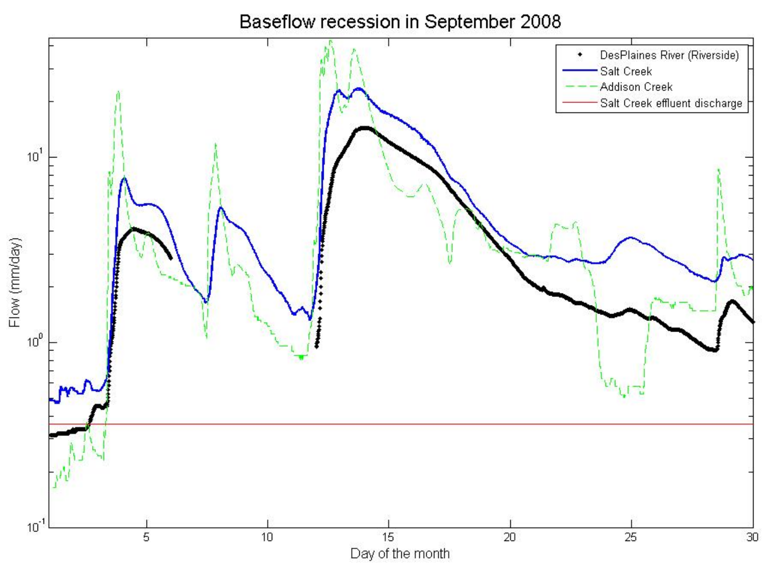

Figure 3.20: Responses of Addison Creek, Salt Creek and downstream Des Plaines River to the September 2008 event, using instantaneous data from USGS. The period from 09/06 to 09/12 is missing for the Des Plaines gage.

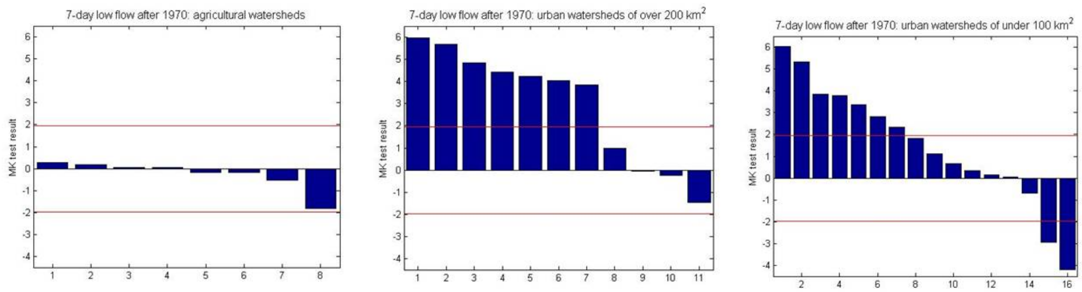

Figure 3.21: For urban areas, we see much more change after 1970 than for peak flow frequency. It is visible at all scales, and neater in older suburbs.

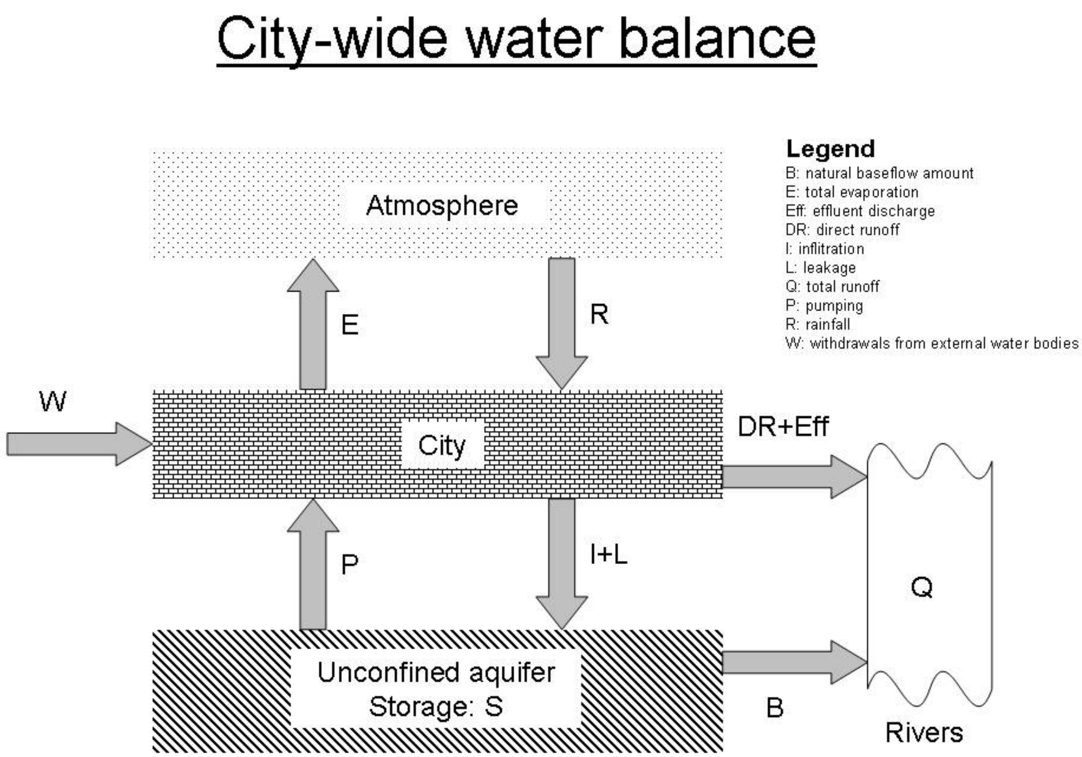

Figure 3.22: Fluxes of water in a city-wide water balance. We assume the unconfined aquifer is local to the city.

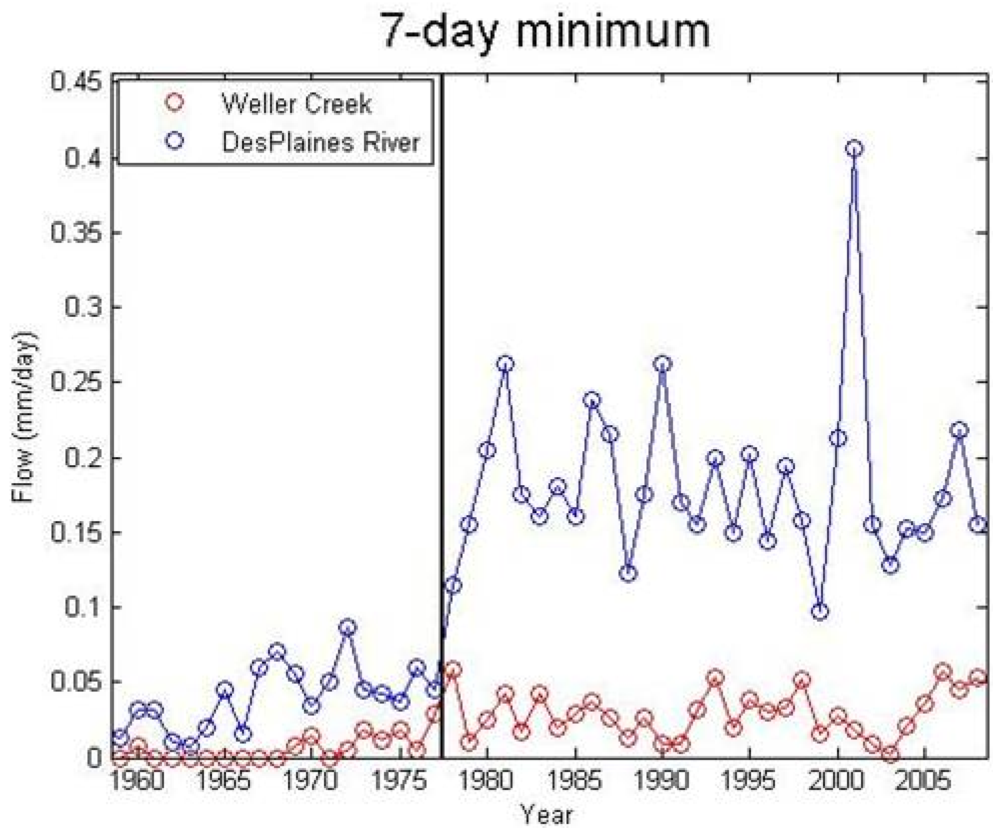

Figure 3.23: This concomitant increase in the large (blue) watershed and its small tributary (red) could be an indication that contrary to a common assumption, low flow are not a local pattern.

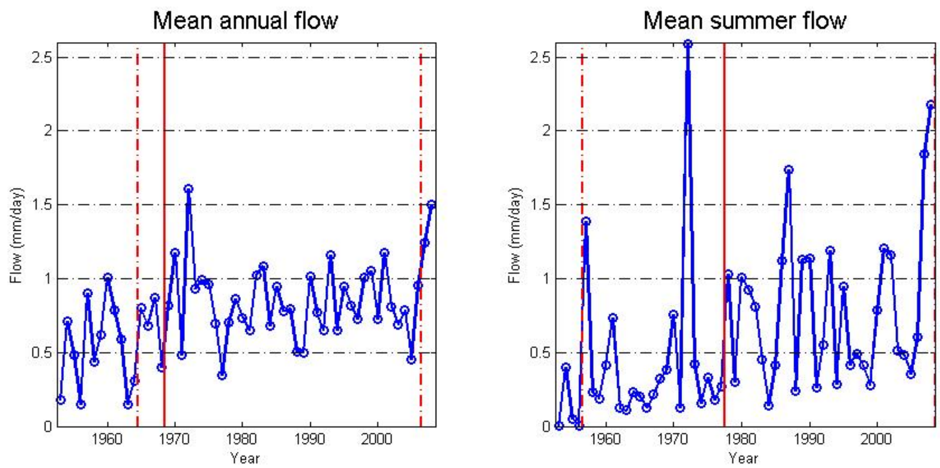

Figure 3.24: For both annual and summer mean discharge, the Pettitt year is the solid red line while the period of change is materialized by the dotted lines. Increases in the low flows and rainfall variability are not suficient to explain the increases we see: this may be the signature of suppressed natural evapotranspiration.

Chapter 4

Conclusions and recommendations

This section summarizes the findings of Chapter 2 and Chapter 3. Section 1 summarizes the development of a method that analyzes the timing, duration and significance of change in a hydrologic time series. Section 2 illustrates how this new method can be applied to the analysis of streamflow non-stationarity at various spatial scales and yield insights on the impacts of climate variability and urban development in the Greater Chicago area. Finally, Section 3 outlines some of the limitations of this work that future studies could tackle.

4.1 Insights on the methodology

The PCPS method is designed to not only determine the significance of a change in a hydrologic time series but also identify the timing and duration of the change. A decomposition of the Mann-Kendall (MK) test statistics is used to identify the location and duration of the period during which the shift of the median occurs. It also has the advantage of using the same rank-based statistics to describe the change pattern and evaluate its significance. This factor, along with its non-parametric nature, makes it a practical framework for understanding the change in a hydrologic time-series. The decomposition is integrated into a rigorous methodological framework that also determines the significance of change through incorporating the recent developments in dealing with the issue of serial correlation with the MK test. Finally, this theoretical framework can also provide data-driven tools to address the important issue of whether that change is likely due to a current source of environmental change or merely a historic one.

Tests of predetermined change patterns demonstrate that the theoretical basis of PCPS enables it to determine the exact dates of change. It is shown that performing an exact detection of the dates of change cannot be done in the presence of noise with rank-based statistical tools alone. As a consequence, PCPS does not decompose the time series into signal and noise components. However, through 10,000 Monte-Carlo simulations, it can be proved that PCPS is both sensitive enough to distinguish between different change patterns, and robust enough to make sure that a large difference between two detected change patterns reflects a large difference between two actual change patterns. This suggests that any change detected over long periods at a set of gages is robust and physically meaningful, which, in turn, demonstrates that this framework is a useful tool for spatial analysis of temporal change patterns.

The applicability of PCPS to spatio-temporal analysis of change has also been justified by the insights that can be discerned from the analysis of hydrologic non-stationarity in northeastern Illinois. Indeed, the application suggests that the method is both sensitive enough to display different change patterns, and robust enough to make it likely that the detection of different change patterns across the chosen dataset actually reflects different causal mechanisms. The use of this method with actual data thus proves that PCPS is a powerful tool for making inferences about relatively large datasets. For instance, when applied to the Greater Chicago area, the causes of change are analyzed for groups of flow indicators that display similar patterns, such as low flows and mean flows. Empirical and conceptual models could also be used to provide a physical interpretation of the observed patterns and, finally, to validate the findings of PCPS.

However, the PCPS framework itself cannot distinguish different causes of streamflow non-stationarity. Natural hydrologic variability confounds the analysis of human interferences on streamflow. This can become especially problematic when non-monotonic patterns, such as low-frequency climatic fluctuations, are present in the data. This problem occurs when the seasonal hydroclimatic data are analyzed in this study. However, it disappears when these fluctuations are translated into an observable shift, as is the case for annual rainfall.

It is possible that with the onset of global climate change, PCPS may become relevant for studying the timing and duration of the shifts it could cause in the water resources throughout the world. The results of the Chicago case study demonstrate that the method is already applicable to city-scale impact studies where the signature of urbanization in the data is robust to climate variability. The findings of the application are presented in the next section.

4.2 City-scale effects of urbanization

The effects of urban development are assessed across different spatial scales within the whole Greater Chicago Area via statistical analysis. First, one has to understand the effects of climatic variability. PCPS shows that a step change in annual precipitation around 1965 coincides with streamflow increases in northeastern Illinois, especially in agricultural areas outside of Chicago. This finding enables to attribute the more gradual increases in streamflow indicators in urban catchments to the interference of other processes. However, seasonal rainfall data do not show as much of a definite pattern, which hinders the empirical analysis of human interferences in streamflow data. Furthermore, local rainfall variability has been shown to make this assessment more complex in some cases.

Thus, we can best identify the effects of urbanization that are statistically robust to the climatic variability displayed in a particular case study. In particular, two major impacts can be found in large areas within the Chicago metropolitan region. They are 1) a reduction and sometimes even an interruption of the increasing trends in peak flows that occur due to the replacement of natural soil storage of stormwater with engineering facilities, such as detention basins, and 2) widespread low and mean flow increases due to water transfers from Lake Michigan for consumptive use at a scale nearing that of the metropolitan area itself. Many of these withdrawals are directly returned to streams through effluent discharges while leakage and garden irrigation also have the potential to augment base flow. These two effects counter two of the direct consequences that occur when infiltration is reduced after making land impervious: increased peak discharges and reduced low flows. In fact, they can be seen as two adaptive responses to two threats that urbanization poses to the area: enhanced and more expensive flooding and an unsustainable water supply due to demographic pressure. Finally, in the Greater Chicago area, the effects of those adaptive responses to urbanization offset those of urbanization itself. The third effect of urbanization, decreased evaporation, reinforces the effects of water withdrawals by gradually increasing mean runoff, especially in the summer.

The temporal patterns induced by these two adaptive responses can be detected across a wide range of scales, but may not appear on the outer edge of the Greater Chicago area, where the effects of urbanization are more diverse and detectable over smaller scales. The adaptive responses described above have not been implemented yet in all of these outer suburbs in which urbanization is typically much less concentrated. In these locations, we can detect positive trends in flood frequency and magnitude, or non-increasing low-flows, with possible withdrawals for on-site water supply. Other recent studies have also thoroughly examined the hydrologic consequences of suburban development (e.g. *Burns et al.*, 2005; *Claessens et al.*, 2006). This suggests that future impact studies should, like this one, take into consideration an urbanized area in its entirety in order to understand which processes dominate at the scales where planning decisions are made. Such studies in other large cities of the world are necessary because climatic, social, economic and demographic conditions vary greatly from city to city. The three later types of factors can determine the pace and spatial extension of urban development, as well as the amount of water required for consumptive use and the resources available for carrying out adaptive responses, such as the ones seen in Chicago.

4.3 Limitations and future work

Even though the non-parametric nature of PCPS is an advantage when dealing with hydrologic data, it is also a limitation because this means that it cannot ascertain the magnitude of change. Although the timing, duration and significance of change are addressed through this method, this shortcoming implies that the insights it provides are only qualitative. As a result, future work could consist of extending PCPS, or coupling it with another tool, to enable it to assess the magnitude of changes as well. This would then allow for the determination of the respective contributions of the different causes that can lead to non-stationarity in a time-series when more than one are present, such as in the example provided with mean flow in this study. Conceptual modeling tools allow for the identification of these causes, but not the quantification of their respective contributions to change.

Other limitations have to do with gaps in the physical understanding of the possible mechanisms of change. They cannot be resolved by further developments of PCPS, because no method can separate the respective contributions of distinct causes if a theoretical basis for understanding these does not exist. For instance, it seems that the possible feedbacks of effluent discharge on groundwater levels is not well-understood. Furthermore, there are no existing base flow separation techniques that take effluent discharges or stormwater management facilities (*Meyer*, 2005). These two limitations, taken together, are an obstacle to quantifying the contribution of low flow increases to mean flow increases.

Another limitation has to do with the quantity of data available to conduct an assessment, such as the one presented in this study. For instance, the lack of accurate and consistently classified historical land-use data tracing back to the beginning of the post-World War II suburban extension prevents observed hydrologic changes from being attributed to land-use change. The same could be said for most historical data related to human activities, such as effluent discharges and groundwater pumping. While this thesis shows the necessity of understanding how socio-economic factors could trigger streamflow change as a side-effect of adaptive responses to issues posed by urbanization, data limitations could make the validation of conceptual models more difficult.

Furthermore, the limited availability of hydrologic data itself limits the range of scales that can be considered in this study, as no gauge collects data in catchments with an area under twenty square kilometers. At smaller spatial scales, the dataset is also heavily biased towards gauges in urbanizing areas, so that no comparison with rural data is available at this scale. Even when facing data availability limitations, innovative thinking can foster a better understanding of non-stationarity as the further development of novel approaches for non-stationary hydrologic time series is critical for addressing hydrologic problems in an era of climate change and rapid urbanization.

References

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005), Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, *438*, doi:10.1038/nature04141.

Barringer, T. H., R. G. Reiser, and C. V. Price (1994), Potential effects of development on flow characterisitics of two new jersey streams, *Water Resources Bulletin*, *30*(2), 283–295.

Bolte, J. P., D. W. Hulse, and S. V. Gregory (2006), Modeling biocomplexity – actors, landscapes and alternative futures, *Environmental Modelling and Software*, *22*, 570–579.

Box, G. E. P., and G. M. Jenkins (Eds.) (1976), *Time-series analysis: forecasting and control*, 575 pp., Holden-Day, San Francisco.

Buishand, T. A. (1982), Some methods for testing the homogeneity of rainfall records, *Journal of Hydrology*, *58*, 11–27.

Burns, D., T. Vitvar, J. McDonnell, J. Hassett, J. Duncan, and C. Kendall (2005), Effects of suburban development on runoff generation in the croton river basin, new york, usa, *Journal of Hydrology*, *311*, 266–281.

Butts, T. A., and D. B. Shackleford (1992), Reduction in peak flows and improvements in water quality in the illinois waterway downstream of lockport due to implementation of phases i and ii of tarp, *Tech. Rep. CR-526*, Illinois State Water Survey, Champaign, IL.

Changnon, S. A., and M. Demissie (1996), Detection of changes in streamflow and floods resulting from climate fluctuations and land use-drainage changes, *Climatic Change*, *32*, 411–421.

Chow, V. T., D. R. Maidment, and L. W. Mays (Eds.) (1988), *Applied Hydrology*, 572 pp., Editions McGraw-Hill, New York.

Claessens, L., C. Hopkinson, E. Rastetter, and J. Vallino (2006), Effect of historical changes in land use and climate on the water budget of an urbanizing watershed, *Water Resources Research*, *42*, W03426, doi:10.1029/2005WR004131.

Cong, Z., D. Yang, B. Gao, H. Yang, and H. Hu (2009), Hydrological trend analysis in the yellow river basin using a distributed hydrological model, *Water Resources Research*, *45*, W00A13, doi: 10.1029/2008WR006852.

Déry, S. J., K. Stahl, R. D. Moore, P. H. Whitfield, B. Menounos, and J. E. Burford (2009), Detection of runoff timing changes in pluvial, nival, and glacial rivers of western canada, *Water Resources Research*, *45*, W04426, doi:10.1029/2008WR006975.

DeWalle, D. R., B. R. Swistock, T. E. Johnson, and K. J. McGuire (2000), Potential effects of climate change and urbanization on mean annual streamflow in the united states, *Water Resources Research*, *36*(9), 2655–2664.

Douglas, E. M., R. M. Vogel, and C. N. Kroll (2000), Trends in floods and low flows in the united states: impact of spatial correlation, *Journal of Hydrology*, *240*, 90–105.

Dow, C. L., and D. R. DeWalle (2000), Trends in evaporation and bowen ratio on urbanizing watersheds in eatern united states, *Water Resources Research*, *36*(7), 1835–1843.

Evans, T. P., and H. Kelley (2004), Multi-scale analysis of a household level agent-based model of landcover change, *Journal of Environmental Management*, *72*, 57–72.

Ferguson, B. K., and P. W. Suckling (1990), Changing rainfall-runoff relationships in the urbanizing peachtree creek watershed, atlanta, georgia, *Water Resources Bulletin*, *26*(2), 313–322.

Fu, G., S. Chen, C. Liu, and D. Shepard (2004), Hydro-climatic trends of the yellow river basin for the last 50 years, *Climatic Change*, *65*, 149–178.

Gerrits, A. M. J., H. H. G. Savenje, E. J. M. Veling, and L. Pfister (2009), Analytical derivation of the budyko curve based on rainfall characteristics and a simple evaporation model, *Water Resources Research*, *45*, W04403, doi:10.1029/2008WR007308.

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou (2002), Advanced spectral methods for climatic time series, *Review of Geophysics*, *40*(1), 1003, doi:10.1029/2000RG000092.

Glantz, M. H. (Ed.) (1999), *Creeping Environmental Problems and Sustainable Development in the Aral Sea Basin*, 291 pp., Cambridge University Press, Cambridge.

Good, P. (Ed.) (2005), *Permutation, parametric and bootstrap tests of hypotheses*, 3rd ed., 315 pp., Springer, New York.

Grimm, N. B., S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, and J. M. Briggs (2002), Global change and the ecology of cities, *Science*, *319*, 756–760.

Grimmond, C. S. B., and T. R. Oke (1986), Urban water balance, 2: Results from a suburb of vancouver, british columbia, *Water Resources Research*, *22*(10), 1404–1412.

Grimmond, C. S. B., and T. R. Oke (1995), Comparison of heat fluxes from summertime observations in the suburbs of four north american cities, *Journal of Applied Meteorology*, *34*, 873–889.

Grimmond, C. S. B., T. R. Oke, and D. G. Steyn (1986), Urban water balance, 1: A model for daily totals, *Water Resources Research*, *22*(10), 1397–1403.

Groisman, P. Y., R. W. Knight, and T. R. Karl (2001), Heavy precipitation and high streamflow in the contiguous united states: Trends in the twentieth century, *Bulletin of the American Meteorological Society*, *82*(2), 219–246.

Groisman, P. Y., R. W. Knight, T. R. Karl, D. R. Easterling, B. Sun, and J. H. Lawrimore (2004), Contemporary changes of the hydrological cycle over the contiguous united states: Trends derived from in situ observations, *Journal of Hydrometeorology*, *5*, 64–85.

Hamed, K. H., and A. R. Rao (1998), A modified mann-kendall trend test for autocorrelated data, *Journal of Hydrology*, *204*, 182–196.

Hejazi, M. I., and M. Markus (2009), Impact of urbanization and climate variability on floods in northeastern illinois, *Journal of Hydrologic Engineering*, *14*(6).

Hejazi, M. I., and G. E. Moglen (2007), Regression-based approach to low flow prediction in the maryland piedmont region under joint climate and land use change, *Hydrological Processes*, *21*, 1793–1801.

Hejazi, M. I., and G. E. Moglen (2008), The effect of climate and land use change on flow duration in the maryland piedmont region, *Hydrological Processes*, *22*, 4710–4722.

Hirsch, R. M., and J. R. Slack (1984), A nonparametric trend test for seasonal data with serial dependence, *Water Resources Research*, *20*(6), 727–732.

Hollis, G. E. (1975), The effect of urbanization on floods of different recurrence intervals, *Water Resources Research*, *11*(3), 431–435.

Hurkmans, R. T. W. L., W. Terink, R. Uijlenhoet, E. J. Moors, P. A. Troch, and P. H. Verburg (2009), Effects of land use changes on streamflow generation in the rhine basin, *Water Resources Research*, *45*, W06405, doi:10.1029/2008WR007574.

Hurst, H. (1951), Long term storage capacities of reservoirs, *Transactions of the American Society of Civil Engineers*, *166*, 776–808.

Karl, T. R., and R. W. Knight (1998), Secular trends of precipitation amount, frequency, and intensity in the united states, *Bulletin of the American Meteorological Society*, *79*(2), 231–241.

Katz, R. W., M. B. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology, *Advances in Water Resources*, *25*, 1287–1304.

Kendall, M. G. (Ed.) (1975), *Rank correlation methods*, 2nd ed., 196 pp., C. Griffin, London.

Koscielny-Bunde, E., J. W. Kantelhardt, P. Braun, A. Bunde, and S. Havlin (2006), Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies, *Journal of Hydrology*, *322*, 120–137.

Koutsoyiannis, D. (2002), The hurst phenomenon and fractional gaussian noise made easy, *Hydrological Sciences Journal*, *47*(4), 573–595.

Koutsoyiannis, D. (2003), Climate change, the hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, *48*(1), 3–24.

Koutsoyiannis, D. (2006), Nonstationarity versus scaling in hydrology, *Journal of Hydrology*, *324*, 239–254.

Koutsoyiannis, D., and A. Montanari (2007), Statistical analysis of hydroclimatic time series: Uncertainty and insights, *Water Resources Research*, *43*, W05429, doi:10.1029/2006WR005592.

Kumar, S., V. Merwade, J. Kam, and K. Thurner (2009), Streamflow trends in indiana: Effects of long term persistence, precipitation and subsurface drains, *Journal of Hydrology*, *374*, 171–183.

Kundzewicz, Z. W., and A. J. Robson (2004), Change detection in hydrological records: a review of the methodology, *Hydrological Sciences Journal*, *49* (1).

Kundzewicz, Z. W., D. Graczyk, T. Maurer, I. Pińskwar, M. Radziejewski, C. Svensson, and M. Szwed (2005), Trend detection in river flow series: 1. annual maximum flow, *Hydrological Sciences Journal*, *50* (5).

Lazaro, T. R. (1976), Nonparametric statistical analysis of annual peak flow data from a recently urbanized watershed, *Water Resources Bulletin*, *12* (1), 101–107.

Lehmann, E. L. (Ed.) (1975), *Nonparametric, statistical methods based on ranks*, Holden-Day, San Francisco.

Lettenmaier, D. P. (1976), Detection of trends in water quality data from records with dependent observations, *Water Resources Research*, *12* (5).

Lettenmaier, D. P., E. F. Wood, and J. R. Walls (1994), Hydro-climatological trends in the continental united states, 1948-1988, *Journal of Climate*, *7*, 586–607.

Lins, H. F., and J. R. Slack (1999), Streamflow trends in the united states, *Geophysical Research Letters*, *26* (2), 227–230.

Lins, H. F., and J. R. Slack (2005), Seasonal and regional characteristics of u.s. streamflow trends in the united states from 1940 to 1999, *Physical Geography*, *26*, 489–501.

Liu, J., T. Dietz, S. Carpenter, C. Folke, M. Alberti, C. Redman, S. Schneider, E. Ostrom, A. Pell, J. Lubchenco, W. Taylor, Z. Ouyang, P. Deadman, T. Kratz, and W. Provencher (2007a), Coupled human and natural systems, *Ambio*, *36* (8).

Liu, J., T. Dietz, S. Carpenter, M. Alberti, C. Folke, E. Moran, A. Pell, P. Deadman, T. Kratz, J. Lubchencoa, E. Ostrom, Z. Ouyang, W. Provencher, C. Redman, S. Schneider, and W. Taylor (2007b), Complexity of coupled human and natural systems, *Science*, *317*.

Liu, P. S. G., L. Xiong, and L. Chen (2010), Flood season segmentation based on probability change point analysis technique, *Hydrologcial Sciences*, in press.

Mandelbrot, B., and J. R. Wallis (1968), Noah, joseph, and operational hydrology, *Water Resources Research*, *4* (5), 909–918.

Mann, H. (1945), Non-parametric tests against trend, *Econometrica*, *13*, 245–259.

Markus, M., J. R. Angel, L. Yang, and M. Hejazi (2007), Changing estimates of design precipitation in northeastern illinois: Comparison between different sources and sensitivity analysis, *Journal of Hydrology*, *347*, 211–222.

Martin, S. (2004), The cost of restoration as a way of defining resilience: a viability approach applied to a model of lake eutrophication, *Ecology and Society*, *9* (2).

McCabe, G., and D. Wolock (2002), A step increase in streamflow in the conterminous united states, *Geophysical Research Letters*, *29* (24), 2185, doi:10.1029/GL015999.

McConkey, S., A. Bartosova, L.-C. Lin, K. Andrew, M. Machesky, and C. Jennings (2004), Fox river watershed investigation - stratton dam to the illinois river: Water quality issues and data report to the fox river study group, inc, *Tech. Rep. CR 2004-2006*, Illinois State Water Survey, Champaign, IL.

McCormick, B. C., K. N. Eshleman, J. L. Griffith, and P. A. Townsend (2009), Detection of flooding responses at the river basin scale enhanced by land-use change, *Water Resources Research*, *45*, W08401, doi:10.1029/2008WR007594.

Meyer, S. C. (2005), Analysis of base flow trends in urban streams, northeastern illinois, usa, *Hydrogeology Journal*, *13*, 871–885.

Milly, P. C. D. (1994a), Climate, soil water storage, and the average annual water balance, *Water Resources Research*, *30*(7), 2143–2156.

Milly, P. C. D. (1994b), Climate, interseasonal storage of soil water, and the annual water balance, *Advances in Water Resources*, *17*, 19–24.

Milly, P. C. D., K. A. Dunne, and A. V. Vecchia (2005), Global pattern of trends in streamflow and water availability in a changing climate, *Nature*, *438*, doi:10.1038/nature04312.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: Whither water management?, *Science*, *319*.

Montanari, A. (2003), Long-range dependence in hydrology, in *Theory and Applications of Long-Range Dependence*, edited by P. D. et al., pp. 461–472, Springer, New York.

Monticino, M., M. Acevedo, B. Callicott, T. Cogdill, and C. Lindquist (2007), Coupled human and natural systems: A multi-agent based approach, *Environmental Modelling and Software*, *22*, 656–663.

Pelletier, J. D., and D. L. Turcotte (1997), Long-range persitence in climatological and hydrological time-series: Analysis, modeling and application to drought hazard assessment, *Journal of Hydrology*, *203*, 198–208.

Perreault, L., M. Haché, M. Slivitzky, and B. Bobée (1999), Detection of changes in precipitation and runoff over eastern canada and u.s. using a bayesian approach, *Stochastic Environmental Research and Risk Assessment*, *13*, 201–216.

Pettitt, A. N. (1979), A non-parametric approach to the change point problem, *Applied Statistics*, *28*(2), 126–135.

Pryor, S. C., J. A. Howe, and K. E. Kunkel (2009), How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous usa?, *International Journal of Climatology*, *29*, 31–45.

Renard, B., M. Lang, and P. Blois (2006), Statistical analyis of extreme events in a non-stationaty context via a bayesian framework: case study with peak-over-threshold data, *Stochastic Environmental Research and Risk Assessment*, *21*, 97–112.

Sankarasubramanian, A., and R. M. Vogel (2003), Hydroclimatology of the continental united states, *Geophysical Research Letters*, *30*(7), 1363, doi:10.1029/2002GL015937.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker (2001), Catastrophic shifts in ecosystems, *Nature*, *413*, 591–596.

Sen, P. K. (1968), Estimates of the regression coefficient based on kendall's tau, *American Statistical Association Journal*, *63*, 1379–1389.

Sivakumar, B. (2008), Nonlinear dynamics and chaos in hydrologic systems: latest developments and a look forward, *Stochastic Environmental Research and Risk Assessment*, *23*(7), 1027–1036.

Small, D., S. Islam, and R. M. Vogel (2006), Trends in precipitation and streamflow in the eastern u.s.: Paradox or perception?, *Geophysical Research Letters*, *33*, L03403, doi: 10.1029/2005GL024995.

Smith, K., and M. B. Richman (1993), Recent hydroclimatic fluctuations and their effects on water resources in illinois, *Climatic Change*, *24*, 249–269.

Solo-Gabriele, H., and F. E. Perkins (1997), Streamflow and suspended sediment transport in an urban environment, *Journal of Hydraulic Engineering*, *123*(9), 807–811.

Svensson, C., Z. W. Kundzewicz, and T. Maurer (2005), Trend detection in river flow series: 2. flood and low-flow index series, *Hydrological Sciences Journal*, *50*(5).

Theil, H. (1950), A rank-invariant method of linear and polynomial regression analysis, i, ii and iii, *Nederl. Akad. Wetensch. Proc.*, *53*, 386–392, 521–525 and 1397–1412.

Tomozeiu, R., A. Busuioc, V. Marletto, F. Zinoni, and C. Cacciamani (2000), Detection of changes in the summer precipitation time series of the region emilia-romagna, italy, *Theoretical and Applied Climatology*, *67*, 193–200.

Vautard, R., P. Yiou, and M. Ghil (1992), Singular-spectrum analysis: a toolkit for short, noisy chaortic signals, *Physica D*, *58*, 95–126.

Villarini, G., F. Serinaldi, J. A. Smith, and W. F. Krajewski (2009a), On the stationarity of annual flood peaks in the continental united states during the 20th century, *Water Resources Research*, *45*, W08417, doi:10.1029/2008WR007645.

Villarini, G., J. A. Smith, F. Serinaldi, J. Bales, P. D. Bates, and W. F. Krajewski (2009b), Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, *Advances in Water Resources*, *32*, 1255–1266.

Vogel, R. (1999), Editorial: Stochastic and deterministic world views, *Journal of Water Resources Planning and Management*, *125*(6), 311–313.

Von Storch, H. (1995), Misuses of statistical analysis in climate research, in *Analysis of Climate Variability: Applications of Statistical Techniques*, edited by N. A. Von Storch H., pp. 11–26, Springer-Verlag, Berlin.

Vörösmarty, C. J., P. Green, J. Salisbury, and R. B. Lammers (2000), Global water resources: Vulnerability from climate change and population growth, *Science*, *289*, 284–288.

Wang, D., and X. Cai (2009), Detecting human inferences to low flows through base flow recession analysis, *Water Resources Research*, *45*, W07426, doi:10.1029/2009WR007819.

Wang, G., J. Xia, and J. Chen (2009), Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: A case study of the chaobai river basin in northern china, *Water Resources Research*, *45*, W00A11, doi:10.1029/2007WR006768.

Werner, B. T., and D. E. McNamara (2007), Dynamics of coupled human-landscape systems, *Geomorphology*, *91*(3-4), 393–407.

Westcott, N. E. (2009), Continued operation of a 25-raingage network for collection, reduction, and analysis of precipitation data for lake michigan diversion accounting: Water year 2008, *Tech. Rep. CR 2009-04*, Illinois State Water Survey, Champaign, IL.

Xiong, L., and S. Guo (2004), Trend test and change-point detection for the annual discharge series of the yangtze river at the yichang hydrological station, *Hydrological Sciences Journal*, *49*(1).

Yeh, C.-H., and J. W. Labadie (1997), Multiobjective watershed-level planning of storm water detention systems, *Journal of Water Resources Planning and Management*, *123*(6), 336–343.

Yue, S., and T. Y. Gan (2004), Simple scaling properties of canadian annual average streamflow, *Advances in Water Resources*, *27*, 481–495.

Yue, S., and C. Y. Wang (2002a), Power of the mann-whitney test for detecting a shift in median or mean of hydro-meteorological data, *Stochastic Environmental Research and Risk Assessment*, *16*, 307–323.

Yue, S., and C. Y. Wang (2002b), The influence of autocorrelation on the mann-whitney test for detecting a shift in median, *Advances in Water Resources*, *25*, 325–333.

Yue, S., and C. Y. Wang (2002c), Applicability of prewhitening to eliminate the influence of serial correlation on the mann-kendall test, *Water Resources Research*, *38*(6), 1068, doi:10.1029/2001WR000861.

Yue, S., P. Pilon, and G. Cavadias (2002a), Power of the mann-kendall and spearman's rho tests for detecting monotonic trends in hydrological series, *Journal of Hydrology*, *259*, 254–271.

Yue, S., P. Pilon, B. Phinney, and G. Cavadias (2002b), The influence of autocorrelation on the ability to detect trend in hydrological series, *Hydrological Processes*, *16*, 1807–1829.

Zehe, E., and M. Sivapalan (2009), Threshold behavior in hydrological systems as (human) geo-ecosystems: manifestations, controls, implications, *Hydrology and Earth System Sciences*, *13*, 1273–1297.

Zhang, Q., C.-Y. Xu, S. Becker, Z. Zhang, Y. Chen, and M. Coulibaly (2009), Trends and abrupt changes if precipitation maxima in the pearl river basin, china, *Atmospheric Science Letters*, *10*, 132–144, doi:10.1002/asi.

Zhang, X., D. Harvey, W. D. Hogg, and T. R. Yuzyk (2001), Trends in canadian streamflow, *Water Resources Research*, *37*(4), 987–998.

Zheng, H., L. Zhang, C. Liu, Q. Shao, and Y. Fukushima (2007), Changes in stream flow regime in headwater catchments of the yellow river basin since the 1950s, *Hydrological Processes*, *21*, 886–893.

Vita

Charles Rougé was born in Toulouse (France) in 1986. He got his bachelor in engineering at the Ecole Polytechnique of Palaiseau (France), before starting his MSc in Civil Engineering at the University of Illinois at Urbana-Champaign in August 2008.